

**UNIVERSIDAD COMPLUTENSE DE MADRID**  
**FACULTAD DE FILOLOGÍA**  
**DEPARTAMENTO DE FILOLOGÍA INGLESA**



**TESIS DOCTORAL**

**Annotating Appraisal in English and Spanish  
product reviews from mobile application stores:  
a contrastive study for linguistic and  
computational purposes**

**La anotación de la Valoración en críticas de productos  
extraídas de tiendas de aplicaciones para móvil en  
inglés y en español: un estudio contrastivo para fines  
lingüísticos y computacionales**

MEMORIA PARA OPTAR AL GRADO DE DOCTORA

PRESENTADA POR

**Natalia Mora López**

DIRECTORA

**Julia Lavid López**

**Madrid, 2018**

# **UNIVERSIDAD COMPLUTENSE DE MADRID**

**FACULTAD DE FILOLOGÍA**  
**Departamento de Filología Inglesa I**  
**(Lengua y Lingüística Inglesa)**



## **TESIS DOCTORAL**

**Annotating Appraisal in English and Spanish product reviews from  
mobile application stores: a contrastive study for linguistic and  
computational purposes**

**La anotación de la Valoración en críticas de productos extraídas de  
tiendas de aplicaciones para móvil en inglés y en español: un estudio  
contrastivo para fines lingüísticos y computacionales**

**MEMORIA PARA OPTAR AL GRADO DE DOCTOR**  
**PRESENTADA POR**

**Natalia Mora López**

**Directora**

**Julia Lavid López**

**Madrid, 2017**

---

**Annotating Appraisal in English and  
Spanish Product Reviews from  
Mobile Application Stores: a  
Contrastive Study for Linguistic and  
Computational Purposes**

---



**DOCTORAL THESIS**

**Natalia Mora López**

Departamento de Filología Inglesa  
Facultad de Filología  
Universidad Complutense de Madrid

Madrid, 2017



La anotación de la Valoración en críticas de  
productos extraídas de tiendas de  
aplicaciones para móvil en inglés y en  
español: un estudio contrastivo para fines  
lingüísticos y computacionales

*Memoria que se presenta para optar al título de Doctor*

**Natalia Mora López**

*Dirigida por*

**Prof. Dra. Dña. Julia Lavid López**

Departamento de Filología Inglesa

Facultad de Filología

Universidad Complutense de Madrid

Madrid, 2017



*To my parents and brother,  
who always try to make  
what is long, short,  
and what is tough, easy.*





## **Acknowledgements**

I would like to acknowledge the effort, insight and exceptional responsiveness of my thesis supervisor, Prof. Dra. Julia Lavid, whose guidance and help has been key in this thesis from the very first day to the very last minute. She always encouraged me not to give up, and her substantial input and thorough revisions made it all possible.

También quiero agradecer a mi familia todo el apoyo y ayuda que me han prestado siempre en todo lo que he hecho, incluido este trabajo, de cualquier forma en que han podido. Gracias.



## **Abstract**

Evaluation, opinion and subjectivity are related phenomena which are currently receiving attention both in the linguistic and the computational communities. One of the most influential theories that deal with the phenomenon of evaluation and subjectivity is the ‘Appraisal’ framework, which proposes that linguistic expressions of evaluative meanings such as emotion, attitude and opinion can be divided into three different axes: Attitude, Engagement and Graduation. To date, there is no large-scale cross-linguistic work in the computational or in the linguistic communities which has set out the task to validate the different features of Appraisal Theory empirically through corpus annotation, and to investigate its application to a relatively new genre, namely, mobile application reviews

The work developed in this dissertation is an attempt at validating aspects of Appraisal Theory in a contrastive manner (i.e. comparing English and Spanish), and at providing a cross-linguistic characterisation of this new review genre in terms of Appraisal features. The categories, empirically validated with a high degree of inter-annotator agreement, are used to annotate a larger bilingual corpus. The results showed interesting language-specific differences, such as a higher degree of straightforward strategies in the Spanish texts in opposition to a higher modulation of meanings in the English ones. In addition, some tendencies may be pointed out among different products (applications and games, vs. books and films) as well as specific aspects that are typically included in negative reviews (Pseudo-Questions and Judgement) or in positive reviews (Capacity).



# Contents

<b>ACKNOWLEDGEMENTS .....</b>	<b>V</b>
<b>ABSTRACT.....</b>	<b>VII</b>
<b>LIST OF FIGURES.....</b>	<b>XI</b>
<b>LIST OF TABLES .....</b>	<b>XIII</b>
<b>CHAPTER 1 INTRODUCTION.....</b>	<b>1</b>
1.1. Background .....	1
1.2. Aims and research questions .....	6
1.3. Methodology.....	7
1.4. Thesis organization .....	8
<b>CHAPTER 2 THEORETICAL FRAMEWORK.....</b>	<b>9</b>
<b>2.1. Appraisal Theory .....</b>	<b>9</b>
2.1.1. Attitude .....	11
2.1.2. Engagement .....	16
2.1.3. Graduation .....	21
2.1.4. Previous and related research on Appraisal .....	24
<b>2.2. The genre of ‘Mobile application reviews’ .....</b>	<b>26</b>
2.2.1. Characteristics of reviews in application stores as a genre.....	29
2.2.2. Previous research on reviews published in application stores .....	35
<b>CHAPTER 3 VALIDATING APPRAISAL FEATURES THROUGH CORPUS ANNOTATION.....</b>	<b>39</b>
<b>3.1. Introduction .....</b>	<b>39</b>
<b>3.2. The ‘training corpus’ .....</b>	<b>40</b>
<b>3.3. Appraisal categories selected for annotation.....</b>	<b>42</b>
3.3.1. Attitude.....	44
3.3.2. Engagement .....	45
3.3.3. Graduation .....	46
<b>3.4. Annotation scheme and guidelines .....</b>	<b>47</b>
<b>3.5. Annotation experiments.....</b>	<b>49</b>
3.5.1. Agreement study 1 .....	49
3.5.2. Agreement study 2 .....	51
3.5.3. Agreement study 3 .....	52

<b>3.6. Evaluating the annotations .....</b>	<b>52</b>
3.6.1. Annotation results for agreement study 1 .....	53
3.6.2. Annotation results for agreement study 2 .....	57
3.6.3. Annotation results for agreement study 3 .....	63
<b>3.7. Summary and discussion.....</b>	<b>67</b>
<b>CHAPTER 4 ANNOTATION OF THE LARGER CORPUS.....</b>	<b>69</b>
<b>4.1. Introduction .....</b>	<b>69</b>
<b>4.2. The larger corpus .....</b>	<b>70</b>
<b>4.3. Annotation results: general tendencies.....</b>	<b>75</b>
<b>4.4. Language-specific differences (English vs. Spanish) .....</b>	<b>77</b>
4.4.1. Attitude .....	79
4.4.2. Engagement.....	83
4.4.3. Graduation .....	88
<b>4.5. Item differences (applications vs. games vs. books vs. films vs. music) .....</b>	<b>91</b>
4.5.1. Attitude .....	92
4.5.2. Engagement.....	96
4.5.3. Graduation .....	99
<b>4.6. Polarity and rating differences .....</b>	<b>101</b>
4.6.1. Attitude .....	102
4.6.2. Engagement.....	105
4.6.3. Graduation .....	108
<b>4.7. Summary and discussion.....</b>	<b>110</b>
<b>CHAPTER 5 CONCLUSION AND FINAL REMARKS.....</b>	<b>115</b>
<b>REFERENCES.....</b>	<b>123</b>
<b>APPENDICES .....</b>	<b>133</b>
<b>Appendix A. CURRICULUM VITAE.....</b>	<b>135</b>
<b>Appendix B. ENGLISH AND SPANISH SUMMARY .....</b>	<b>139</b>
<b>Appendix C. QUANTIFICATION OF TOTAL COMMENTS EXTRACTED.....</b>	<b>151</b>
<b>Appendix D. ANNOTATION GUIDELINES.....</b>	<b>155</b>
<b>Appendix E. TRAINING CORPUS .....</b>	<b>165</b>
<b>Appendix F. LARGER CORPUS .....</b>	<b>175</b>
<b>Appendix G. ANNOTATED CORPORA .....</b>	<b>213</b>

# List of Figures

FIGURE 1. APPRAISAL SYSTEM (ADAPTED FROM MARTIN AND WHITE 2005: 38) .....	10
FIGURE 2. ATTITUDE SYSTEM.....	11
FIGURE 3. THE ENGAGEMENT SYSTEM (ADAPTED FROM MARTIN AND WHITE 2005: 134).....	17
FIGURE 4. THE GRADUATION SYSTEM .....	22
FIGURE 5. STANDARD REVIEW DISPLAY (WHATSAPP, PC VERSION) .....	30
FIGURE 6. STANDARD REVIEW WRITING DISPLAY .....	30
FIGURE 7. COMMENTS RATING BY OTHER USERS.....	32
FIGURE 8. STAR RATING MEANING .....	33
FIGURE 9. LOW RATING DUE TO LOW TECHNOLOGICAL EXPERTISE .....	34
FIGURE 10. USERS HELP OTHER USERS .....	35
FIGURE 11. APPRAISAL CATEGORIES SELECTED FOR ANNOTATION.....	43
FIGURE 12. DISTRIBUTION OF AGREEMENT AND DISAGREEMENT IN MARKABLE IDENTIFICATION .....	55
FIGURE 13. DISAGREEMENT DISTRIBUTION OF MAJOR CATEGORIES .....	60
FIGURE 14. DISAGREEMENT DISTRIBUTION OF MAJOR SUBCATEGORIES .....	61
FIGURE 15. DISAGREEMENT DISTRIBUTION OF CATEGORIES COMBINATION.....	62
FIGURE 16. DISAGREEMENT DISTRIBUTION OF MINOR CATEGORIES .....	67
FIGURE 17. UAM CORPUS TOOL HOME SCREEN .....	72
FIGURE 18. UAM CORPUS TOOL ANNOTATION SCREEN .....	72
FIGURE 19. DISTRIBUTION OF OCCURRENCES PER CATEGORY IN THE LARGER DATASET .....	75
FIGURE 20. DISTRIBUTION OF APPRAISAL FEATURES IN ENGLISH AND SPANISH DATASETS .....	78
FIGURE 21. DISTRIBUTION OF ATTITUDE IN ENGLISH AND SPANISH .....	80
FIGURE 22. DISTRIBUTION OF AFFECT IN ENGLISH AND SPANISH .....	80
FIGURE 23. DISTRIBUTION OF JUDGEMENT IN ENGLISH AND SPANISH .....	81
FIGURE 24. DISTRIBUTION OF APPRECIATION IN ENGLISH AND SPANISH .....	82
FIGURE 25. DISTRIBUTION OF ENGAGEMENT IN ENGLISH AND SPANISH .....	84
FIGURE 26. DISTRIBUTION OF EXPANSION IN ENGLISH AND SPANISH .....	85
FIGURE 27. DISTRIBUTION OF CONTRACTION IN ENGLISH AND SPANISH .....	86
FIGURE 28. DISTRIBUTION OF GRADUATION IN ENGLISH AND SPANISH .....	88
FIGURE 29. DISTRIBUTION OF FORCE AND FOCUS IN ENGLISH AND SPANISH.....	89
FIGURE 30. DISTRIBUTION OF APPRAISAL IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC.....	91
FIGURE 31. DISTRIBUTION OF ATTITUDE IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC .....	92
FIGURE 32. DISTRIBUTION OF AFFECT IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC .....	93
FIGURE 33. DISTRIBUTION OF JUDGEMENT IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC.....	94
FIGURE 34. DISTRIBUTION OF APPRECIATION IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC .....	95
FIGURE 35. DISTRIBUTION OF ENGAGEMENT IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC .....	96

FIGURE 36 DISTRIBUTION OF EXPANSION IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC.....	97
FIGURE 37. DISTRIBUTION OF CONTRACTION IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC.....	98
FIGURE 38. DISTRIBUTION OF GRADUATION IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC.....	99
FIGURE 39. DISTRIBUTION OF FORCE AND FOCUS IN APPLICATIONS, GAMES, BOOKS, FILMS AND MUSIC .....	100
FIGURE 40. DISTRIBUTION OF APPRAISAL IN POSITIVE AND NEGATIVE REVIEWS .....	102
FIGURE 41. DISTRIBUTION OF ATTITUDE IN POSITIVE AND NEGATIVE REVIEWS .....	102
FIGURE 42. DISTRIBUTION OF AFFECT IN POSITIVE AND NEGATIVE REVIEWS.....	103
FIGURE 43. DISTRIBUTION OF JUDGEMENT IN POSITIVE AND NEGATIVE REVIEWS .....	104
FIGURE 44. DISTRIBUTION OF APPRECIATION IN POSITIVE AND NEGATIVE REVIEWS .....	105
FIGURE 45. DISTRIBUTION OF ENGAGEMENT IN POSITIVE AND NEGATIVE REVIEWS.....	106
FIGURE 46. DISTRIBUTION OF EXPANSION IN POSITIVE AND NEGATIVE REVIEWS .....	106
FIGURE 47. DISTRIBUTION OF CONTRACTION IN POSITIVE AND NEGATIVE REVIEWS .....	107
FIGURE 48. DISTRIBUTION OF GRADUATION IN POSITIVE AND NEGATIVE REVIEWS.....	108
FIGURE 49. DISTRIBUTION OF FORCE AND FOCUS IN POSITIVE AND NEGATIVE REVIEWS .....	109



# List of Tables

TABLE 1. SELECTION OF 25 ENGLISH TEXTS FOR THE TRAINING CORPUS .....	41
TABLE 2. SELECTION OF 25 SPANISH TEXTS FOR THE TRAINING CORPUS .....	41
TABLE 3. SUMMARY OF ENGLISH AND SPANISH COMMENTS EXTRACTED .....	42
TABLE 4. CORE TAGSET FOR APPRAISAL ANALYSIS .....	47
TABLE 5. EXTENDED TAGSET FOR APPRAISAL ANNOTATION .....	48
TABLE 6. INTERPRETATION OF <i>KAPPA</i> (FROM VIERA AND GARRET 2005: 362) .....	53
TABLE 7. INTER-ANNOTATOR AGREEMENT IN SPAN SELECTION .....	54
TABLE 8. INTER-ANNOTATOR AGREEMENT IN CORE TAGSET SELECTION .....	58
TABLE 9. INTER-ANNOTATOR AGREEMENT IN MINOR ATTITUDE CATEGORIES SELECTION .....	63
TABLE 10. INTER-ANNOTATOR AGREEMENT IN MINOR ENGAGEMENT CATEGORIES SELECTION .....	65
TABLE 11. INTER-ANNOTATOR AGREEMENT IN MINOR GRADUATION CATEGORIES SELECTION .....	66
TABLE 12. SELECTION OF 100 ENGLISH TEXTS FOR THE LARGER DATASET .....	70
TABLE 13. SELECTION OF 100 SPANISH TEXTS FOR THE LARGER DATASET .....	70
TABLE 14. CORE TAGSET FOR APPRAISAL ANNOTATION .....	73
TABLE 15. EXTENDED TAGSET FOR APPRAISAL ANNOTATION .....	74
TABLE 16. MOST COMMON APPRAISAL SPANS IN ATTITUDE, ENGAGEMENT AND GRADUATION .....	77
TABLE 17. MOST COMMON APPRAISAL SPANS IN ENGLISH AND SPANISH .....	79
TABLE 18. MOST COMMON ATTITUDE SPANS IN ENGLISH AND SPANISH .....	83
TABLE 19. MOST COMMON ENGAGEMENT SPANS IN ENGLISH AND SPANISH .....	87
TABLE 20. MOST COMMON GRADUATION SPANS IN ENGLISH AND SPANISH .....	90
TABLE C1. TOTAL NUMBER OF APPLICATION COMMENTS EXTRACTED .....	151
TABLE C2. TOTAL NUMBER OF GAME COMMENTS EXTRACTED .....	151
TABLE C3. TOTAL NUMBER OF BOOK COMMENTS EXTRACTED .....	152
TABLE C4. TOTAL NUMBER OF FILM COMMENTS EXTRACTED .....	152
TABLE C5. TOTAL NUMBER OF MUSIC COMMENTS EXTRACTED .....	153

---

# Chapter 1

## INTRODUCTION

### 1.1. Background

Evaluation, opinion and subjectivity are related phenomena which are currently receiving attention both in the linguistic and the computational communities under different theories and approaches.

In the computational community, there is a growing body of research in what is known as ‘subjectivity analysis’ (Wiebe et al. 2004) and ‘sentiment analysis’ (Pang et al. 2002; Taboada 2009; Taboada and Grieve 2004; Whitelaw et al. 2005; Bloom, Garg and Argamon 2007; Bloom, Stein and Argamon 2007; Pang and Lee 2008, *inter alia*) as a result of the explosion of opinion expressed by people in Web 2.0 applications such as personal websites, blogs, social networking sites, forums and review sites. The opinions expressed in these sites are now used by organizations to know how the public has received their product instead of conducting surveys, and this information is “a valuable resource for marketing intelligence, social psychologists and others interested in extracting and mining views, moods and attitude” (Kumar and Sebastian 2012: 1).

These studies focus basically on English but there have also been recent extensions to other languages such as Chinese (Tan and Zhang 2008), Italian (Di Bari,

Sharoff and Thomas 2015), Arabic (Abdulla et al. 2016) and Spanish (Moreno Ortiz et al. 2010b, Moreno Ortiz et al. 2011, Boldrini et al. 2009, Cruz et al. 2008, Balahur et al., 2009, inter alia)<sup>1</sup>.

The perspectives addressed have been varied: Prabowo and Thelwall (2009) mix a rule-based classification, supervised learning and machine learning into a new combined method. Kontopoulos et al. (2013) propose the deployment of original ontology-based techniques towards a more efficient sentiment analysis of Twitter posts, by means of receiving a sentiment grade for each distinct notion in the post. Poria et al. (2014) introduce a novel paradigm to concept-level sentiment analysis that merges linguistics, common-sense computing, and machine learning for improving the accuracy of tasks such as polarity detection. Tang, Qin and Liu (2015) provide an overview of the successful deep learning approaches for sentiment analysis tasks.

Also, different linguistic levels have been studied. Titov and McDonald (2008) propose a statistical model to assign words to a set of induced topics, in order to improve aspect-based sentiment summarisation. Turney (2002) addressed the phrase level by applying an unsupervised learning algorithm for classifying reviews as recommended or not recommended to 410 reviews from Epinions. Wilson, Wiebe and Hoffmann (2009) also address the phrase-level focusing on the automatic distinction between prior polarity (or semantic orientation) and contextual polarity (the polarity of the phrase in use). Yu and Hatzivassiloglou (2003) work on newspapers texts and identify a positive, neutral or negative polarity at the sentence level. Taboada, Brooke and Stede (2009) present a taxonomy and classification system for distinguishing between different types of paragraphs in movie reviews: formal vs. functional paragraphs and, within the latter, between description and comment.

Finally, different types of texts have also been investigated. For example, Li and Wu (2010) employ online forums for the sentiment analysis. Schumaker et al. (2012) evaluate sentiment on financial texts. Konerekand and Šimko (2013) introduce a method to recognize opinions in microblog posts. Ortigosa et al. (2014) base their study on Facebook messages written by users.

All these computational studies require the development of machine-readable annotated corpora with specific features reflecting opinion and subjectivity which can

---

<sup>1</sup> For a thorough review of the area of Sentiment Analysis from a linguistic perspective, see Taboada (2016).

later be used as test data for automating the annotation process. The creation of such annotated corpora, if enriched with high-level semantic, pragmatic and discourse features, can be a laborious and time-consuming task.<sup>2</sup>

In spite of the challenges involved in this area, recent years have witnessed the development of important annotated resources for English such as the Sentiment Treebank (Socher et al. 2013) or the MPQA opinion corpus (Wiebe et al. 2005; Wilson 2007). The Sentiment Treebank includes annotations on film reviews mixing the degree of the annotations from very small spans to larger segments of sentences. This means that usual obstacles like the order of words or information beyond the assignation of positive and negative points to positive and negative words are overcome, since this resource is built up on sentence structure. Moreover, it is presented as the only model that captures not only negation but also contrastive conjunctives most effectively.

The MPQA corpus provides an infrastructure for sentiment annotation that cannot be found in other sentiment corpora intended for NLP. It consists of a collection of news articles from a variety of sources that are manually annotated according to opinions and private states such as beliefs or emotions. More recently, Deng and Wiebe (2015) have enriched the MPQA corpus with entity and event target annotations, that is, the targets of particular events are specified and given that role. This allows identifying to whom negative or positive comments are addressed and obtaining a more accurate value of the opinion conveyed in the text.

An important strand of research has focused on the annotation of reviews of consumer products, ranging from online products such as digital cameras (Hu and Liu 2004; Titov and McDonald 2008) to reviews of films, books and hotels (Taboada 2008). Other researchers have annotated non-review texts with semantic or syntactic roles. This is the case of Deng et al. (2013), who used the corpus collected in Conrad et al. (2012), consisting of online editorials and blog posts concerning the debate over the recent health insurance reform in the United States. Deng et al. enrich this corpus with information regarding how entities are affected either positively or negatively by events and the writer's attitude towards their agents and objects. This semantic information beyond positive and negative polarity offer additional details to raise accuracy in automatic sentiment analysis.

---

<sup>2</sup>See Lavid (2017) for a thorough description of the issues and challenges involved in the annotation of high-level discourse features in English and Spanish texts.

In the Spanish context, the work of Moreno-Ortiz and co-authors deserves special attention as it clearly stands out in the field of Sentiment Analysis in Spain. Their work has addressed different domains (financial texts, running shoes reviews, hotel reviews, Twitter messages, etc.) creating specialised corpora for the development of sentiment analysis systems (see Moreno-Ortiz and Gómez Pascual 2016; Moreno-Ortiz and Fernández Cruz 2015; Moreno-Ortiz and Pérez Hernández 2013; Moreno-Ortiz and Pineda Castillo and Hidalgo García 2010) and also extending their work to multiple-domain opinion texts (Moreno-Ortiz et al. 2011a, 2011b). Their work is lexically-based, not discourse-oriented as the current one, but deserves a special mention due to its impact and relevance in the Spanish NLP community, in which remarkable studies are also carried out (see Vilares et al. 2015; Molina-González et al. 2013).

In the linguistic community, corpus-based research on evaluation and subjectivity in language has flourished from different theoretical and applied perspectives. Important theoretical approaches have been provided by researchers such as Hunston (2003, 2010), Hunston and Thompson (2000) on English, and more recent studies including other European languages (see Thompson and Alba-Juez 2014). Hunston and Thompson (2000: 5) define evaluation as

the broad cover term for the expression of the speaker or writer's attitude or stance towards, a viewpoint on, or feelings about the entities or propositions that he or she is talking about...

and they establish that evaluation is conceptually comparative, subjective and value laden. Moreover, they provide four parameters of evaluation, namely: good-bad, certainty, expectedness and importance, being the first one the most important one (Hunston and Thompson 2000: 25). In addition, Thompson and Alba-Juez (2014:6-9) indicate that evaluation does not only affect lexis, but it permeates all language levels, either explicitly or implicitly, and involves a discursive process rather than a textual phenomenon.

One of the most influential theories that deals with the phenomenon of evaluation and subjectivity is the 'Appraisal' framework, a linguistic approach for 'exploring, describing and explaining the way language is used to evaluate, to adopt stances, to construct textual personas and to manage interpersonal positionings and relationships' (see Appraisal Homepage at <http://grammatics.com/appraisal/>). Appraisal Theory has

been developed by Martin and White in Australia (see Martin 2000; Martin and White 2005; White 2003; White 2004, *inter alia*) and proposes that linguistic expressions of evaluative meanings such as emotion, attitude and opinion can be divided into three different axes: Attitude, Engagement and Graduation. A detailed review of the theory is presented in chapter 2.

As a result of the great impact of this theory, considerable research has been devoted to applying and expanding the theory in different ways, focusing mainly on the area of Attitude (e.g. Macken-Horarik 2003; Painter 2003; Page 2003; White 2004), with less attention being paid to the areas of Engagement and Graduation, which require further study. The work developed has basically focused on English, although some cross-linguistic studies involving both European and non-European languages has emerged during the last decade. This includes contrastive work between English and Spanish journalistic texts (Marín and Perucha 2006; McCabe 2007), consumer reviews (Mora 2011, Carretero and Taboada 2009, 2010a, 2010b, 2011, 2014) and other text types (Taboada, Carretero and Hinnel, 2014; Lavid et al. 2014; Lavid, Carretero and Zamorano 2016).

Appraisal Theory has also been applied to other languages and to different genres. Some examples include the work on Japanese on blog articles (Sano 2011), on Brazilian Portuguese in a narrative text (Vian Jr. 2008), on Italian letters to the editor (Pounds 2010), on German political interviews (Becker 2009). In addition, comparative work on different languages has also been carried out on English, French and Indonesian on news reports (Thomson, White and Kitley 2008), or more recent work comparing Appraisal in Spanish, English and German in books and film reviews (Taboada, Carretero and Hinnel 2014).

However, to date, there is no large-scale cross-linguistic work in the computational or in the linguistic communities which has set out the task to validate the different features of Appraisal Theory empirically through corpus annotation, and to investigate its application to a relatively new genre, namely, mobile application reviews.

The work developed in this dissertation is an attempt at validating aspects of Appraisal Theory in a contrastive manner (i.e. comparing English and Spanish), and at providing a cross-linguistic characterisation of this new review genre in terms of Appraisal features.

My initial hypothesis is that mobile application reviews are characterised by specific Appraisal features which are a reflection of the discourse communities where they are produced. In other words, the type and distribution of Appraisal features in this genre is not random but is determined by generic variables such as the communicative purpose of these texts and the language-specific style of the discourse community where they are produced.

### **1.2. Aims and research questions**

The present study has several aims, each of them oriented to the research fields where the current dissertation is framed: Corpus Linguistics, Natural Language Processing (NLP) and Contrastive functional Linguistics.

From the Corpus Linguistic and NLP perspectives, the main aims are:

- a) to compile a bilingual (English-Spanish) textual database of English and Spanish opinion texts, which abound in expressions of Appraisal. More specifically, comments published at Play Store about mobile applications and games, music, books and films.
- b) to validate aspects of the Appraisal Theory empirically and cross-linguistically using corpus annotation and agreement studies as tools for such a purpose.
- c) to enrich the texts with Appraisal features, in order to create a quality-annotated resource which can be used for a number of linguistic and computational purposes.

From the Contrastive Functional perspective, the main aim is to provide an empirically-tested and contrastive characterisation of the selected texts in terms of Appraisal features, which allow the contrastive investigation of the existing variation in the English and the Spanish comments.

The research questions posed and investigated in this study are the following ones:

- (1) Is it possible to test the validity of aspects of Appraisal Theory empirically through corpus annotation? If so, are there features which need redefinition or reformulation to ensure high-quality bilingual (English-Spanish) annotation?



- (2) Are there language-independent Appraisal features which characterise mobile application reviews in English and Spanish? If so, which are these features?
- (3) Are there language-specific Appraisal features of mobile application reviews written in English and Spanish?
- (4) Are there significant differences in the use of Appraisal features depending on the type of item reviewed and depending on the polarity features of the review?

### **1.3. Methodology**

In order to investigate these research questions described above, the following methodological steps were carried out:

1. In order to address research question (1), i.e., the empirical validation of Appraisal Theory through corpus annotation, the methodology put forward by Lavid (2012) and Hovy and Lavid (2010) was applied. This consisted of the following phases:
  - a. A bilingual database of Spanish and English opinion texts was compiled from thousands of comments in Google's Play Store using a crawler designed *ad hoc*. The extracted texts were filtered and selected according to different criteria: language, length, polarity, topic and rating.
  - b. An annotation scheme and guidelines were designed on the basis of the main features proposed in Appraisal Theory, along the three axes of Attitude, Engagement and Graduation.
  - c. Annotation experiments were designed to test the empirical validity of the annotation scheme. These were carried out by two independent annotators working separately on a training corpus of fifty texts extracted randomly from the larger bilingual dataset.
2. In order to address research question (2), a larger corpus was single-annotated by the author of this dissertation with the validated Appraisal tags of the annotation scheme developed in 1.b.
3. In order to address research questions (3) and (4), the distribution of Appraisal tags was examined in the English and the Spanish datasets, considering the type of item reviewed and the polarity features of the review, i.e., positive or negative.

## 1.4. Thesis organization

This study is structured as follows:

The current chapter (numbered as Chapter 1 for convenience) offers an introduction to the study, including the background in which the study is framed, the aims and research questions that have been investigated, the methodology followed in the experiments and tasks and the thesis organisation.

Chapter 2 provides an overview of the theoretical concepts used throughout the study. This include: a) an in-depth description of Appraisal Theory, as proposed within Systemic Functional Linguistics (SFL); b) a characterisation of the genre of ‘mobile application reviews’.

Chapter 3 describes the training corpus selected and the three agreement experiments carried out to validate aspects of Appraisal Theory empirically and cross-linguistically, including the design of annotation guidelines, the procedure of the agreement studies and the discussion of the annotation results.

Chapter 4 is devoted to the presentation of the annotation results of the larger database of bilingual texts. The data are arranged according to language differences (Spanish and English), item differences (applications, games, books, films and music) and polarity differences (positive and negative reviews).

Chapter 5 discusses the main findings and answers the research questions posed in Chapter 1, as well as outlining some further research that would enrich the present work.

## Chapter 2

# THEORETICAL FRAMEWORK

This section provides an overview of the two main areas and concepts on which this study is based. These are the notions of the genre of ‘reviews’ and the systemic-functional linguistic theory of “Appraisal” and its subtypes.

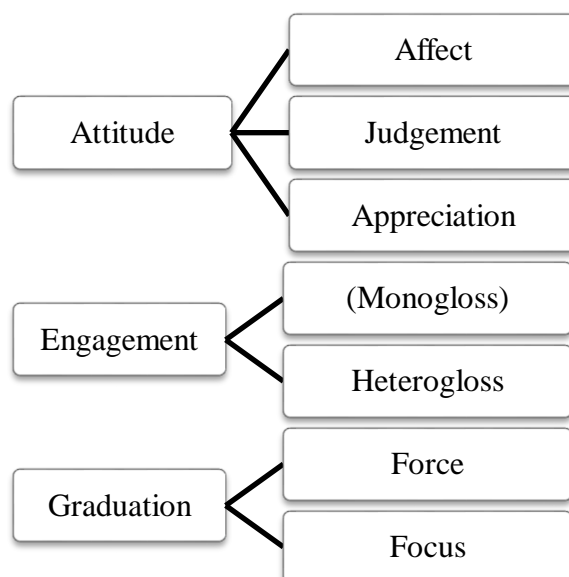
### **2.1. Appraisal Theory**

Appraisal Theory has been mainly developed in Australia, as their most important contributors are J.R. Martin and P. White (Martin 2000; White 2003; Martin and White 2005). This theory is widely explained in Martin and White’s book *The Language of Evaluation*, published in 2005, although some previous works can be found in which the theory was initially suggested and delineated (Martin 2000; White 2002). Martin and White’s research lines show their interests in the application of their theory to, mainly but not only, the journalistic and educational fields (Martin 2001; White 1998, 2004, 2011).

As explained above, Martin and White’s model of evaluation emerged within the theoretical framework of SFL. More specifically, they describe it as “an interpersonal system at the level of discourse semantics” (Martin and White 2005: 33). It is an interpersonal system as far as it evolves within the interpersonal metafunction (i.e. the

negotiation of social relationships), and discourse semantics refers to the meaning beyond the clause. Therefore, Appraisal deals with interpersonal meaning above the clause level.

Appraisal is realised in the form of evaluative language, which is created by expressions that indicate “the subjective presence of writers/speakers in texts as they adopt stances towards both the material they present and those with whom they communicate” (Martin and White 2005: 1). As Hunston (2010: 20) points out, even though most of examples used by these authors are lexical items, not only isolated words convey these meanings<sup>3</sup>. Following the systemic functional tenets, appraisal resources are considered as a system of their own within language, and can be divided into three subsystems: Attitude, Graduation and Engagement, presented in Figure 1.



**Figure 1. Appraisal system (adapted from Martin and White 2005: 38)**

Attitude is concerned mainly with feelings, such as emotions, judgements and evaluations; Engagement is concerned with the ways in which the speaker or writer positions themselves towards the text and other possible voices; and finally, Graduation is concerned with the degrees of intensity of the meanings expressed by Attitude and Engagement realisations. In the following subsections, each of these three axes and their specific types are explained in more detail.

---

<sup>3</sup> Actually, some attempts have been made to classify the occurrence of Attitude (Affect, Judgement and Appreciation) into fixed patterns (Martin 2003, Taboada and Grieve 2004).

### 2.1.1. Attitude

Attitude is used to communicate three types of meaning: emotional (Affect), ethical (Judgement) and aesthetical (Appreciation) (Martin and White 2005: 42). This system may be considered as the most important one for evaluative language, since it is linked to the prototypical opinionated items (typically adjectives and adverbs). In a similar way, following Painter (2003), they claim that, considering Affect, Judgement and Appreciation, the three subsystems of Attitude, it is the first one the most central one, inasmuch as we are born with and embody physiologically this expressive resource. Martin and Rose (2003: 24) summarize these three categories in the following clear and brief way: Affect refers to “resources for expressing feelings”, Judgement to “resources for judging character” and Appreciation to “resources for valuing the worth of things”. Figure 2 summarises this system.

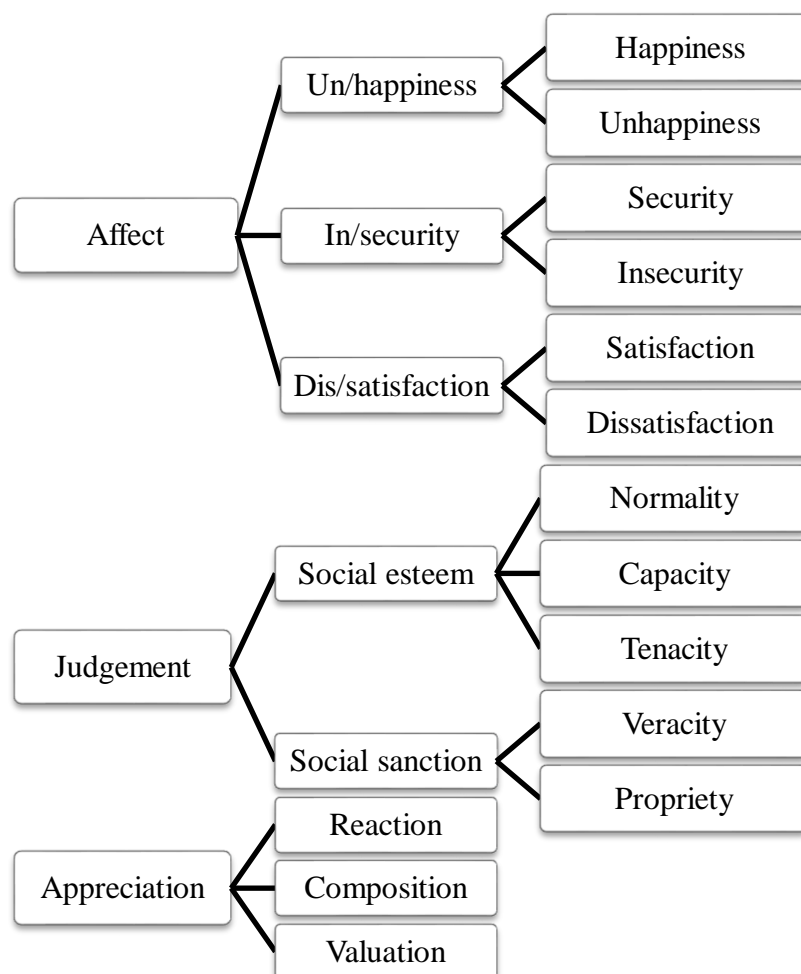


Figure 2. Attitude system

### 2.1.1.1. Affect

Generally speaking, Affect expresses positive or negative feelings and emotional reactions, as mentioned above. Some of the emotional responses it may convey are happiness, anger, disgust, etc. about someone's reactions, either the speaker or a different person. Affection assessments are usually realised in three ways (Martin and White 2005: 46): as qualities (adjectives and adverbs), as processes (verbs) and as comments (sentence adjuncts).

Affect is divided into three sets that cover some groups of emotions. These sets are happiness and unhappiness, security and insecurity, and satisfaction and dissatisfaction.

#### a) Un/happiness

Happiness and unhappiness cover positive feelings of cheer and affection, such as *happy*, *laugh*, *jubilant*, *like*, etc. and negative meanings of misery and antipathy such as *sad*, *cry*, *dislike*, *hate*, etc. Some examples of actual usage are shown in [1]-[4].

- [1] I really hated Avatar (61)<sup>4</sup>
- [2] I'm glad I didn't read anymore (50)
- [3] Me encanta Adele [I love Adele] (191)
- [4] Una lástima porque perdes e l interés [A pity because you lose the interest] (127)

#### b) In/security

Security and insecurity include moods related to anxiety, disquiet and surprise, such as *restless*, *uneasy*, *freaked out*, etc., and confidence and trust, such as *entrust*, *comfortable with*, *confident in/about*, etc. [5]-[8] illustrate this set.

---

<sup>4</sup> Examples have been taken from the corpus analysed, therefore, they may include orthographic mistakes to reflect the way they are originally written. Numbers like (1) belong to the larger corpus, (11) and the like belong to the training corpus. Otherwise, examples have been taken or adapted from Martin and White (2005).

- [5] I was pleasantly surprised by some tracks (92)
- [6] I am afraid to even use this app anymore (2)
- [7] Me asombró el final (157) [The end surprised me]
- [8] Me da miedo dejárselo a mi hija (132) [I am afraid of lending it to my daughter]

c) Dis/satisfaction

Finally, satisfaction and dissatisfaction include meanings involving interest and pleasure, such as *absorbed* or *pleased*, and disappointment and similar feelings of ennui and displeasure, such as *bored with*, *angry*, *fed up with*, etc. A couple of instances are provided in [9]-[12].

- [9] I was disappointed in “25” (87)
- [10] You’ll enjoy it (42)
- [11] Lo único que molesta un poco son todas esas utilidades extras (106) [The only thing that bothers me a bit are all those extra utilities]
- [12] Espero que solucionen esto lo antes posible (111) [I hope that you solve this as soon as possible]

### 2.1.1.2. Judgement

Judgement assessments provide moral evaluations or attitudes about people and the way they behave according to normative rules not established by the individual, either the speakers themselves or someone the speaker is talking about, as Affect is, but by society’s perception of esteem and sanction.

a) Social esteem

Social esteem, that is, those behaviours which are approved and appreciated by society, have to do with three aspects: Normality, Capacity and Tenacity. The first one includes assessments that value how usual or unusual someone or their behaviour is. The second

one answers the question of how capable someone is regarding some aspect or action. The third one conveys meanings of resolution and bravery and their opposites.

All these three aspects can be addressed either from a positive or a negative point of view. Some examples of Normality ([13]-[14]), Capacity ([15]-[16]) and Tenacity ([17]-[18]) can be found in the following ones.

- [13] It took me entirely into the somewhat atypical worlds (47)
- [14] Está de moda (115) [It's fashionable]
- [15] She's still a talented singer (91)
- [16] Solo sirve con dispositivos lentos (109) [It only works with slow devices]
- [17] It has become invasive (2)
- [18] Tengan cuidado (112) [Be careful]

### b) Social sanction

Although social sanction may sound like the negative counterpart of social esteem, it actually deals with very different meanings. On the one hand, it includes assessments about Veracity, i.e., how honest or truthful someone can be considered. On the other hand, it also covers meanings about Propriety, where one can find values more ethically-oriented. Instances [19]-[20], for the case of Veracity, and [21]-[22], for Propriety cases, illustrate this type of Judgement assessments.

- [19] It's a virus or fraud (7)
- [20] Muy poco creíble (145) [Very little credible]
- [21] For his own greed (69)
- [22] Me parece legítimo (132) [It seems legitimate to me]

### 2.1.1.3. Appreciation

While Affect was directly related to the speaker and Judgement to how the speaker judges other people and their behaviour, Appreciation turns its attention to things. This attention is divided into three points of view: the way the speaker reacts to these things, what these things are made of, and what their value is.



## a) Reaction

Reaction assessments include all those expressions that communicate the reaction an item produces or may produce in the speaker or according to the speaker's opinion (either considering whether it caught speakers attention or whether they liked or disliked it). [23]-[24] illustrate these meanings.

[23] This is amazing (84)

[24] She has a beautiful voice (96)

[25] Y más emocionantes (177) [And more exciting]

[26] Me ha parecido un libro diferente que te atrapa (143) [I seemed to me a very different book that captivates you]

## b) Composition

Composition meanings are conveyed in assessments where the complexity and balance of the item are expressed. This is exemplified in [27]-[28].

[27] Intricate, balanced, brilliant story (48)

[28] Se hace fácil de leer (141) [It turns easy to read]

## c) Valuation

The last type of appreciation, Valuation, involves words and expressions where something is valued as worthy or unworthy. These meanings are shown in examples [29]-[32].

[29] They made a landmark rock album (82)

[30] What a waste of money (81)

[31] Muy sobrevalorada (168) [Very overrated]

[32] Es una nueva obra maestra (172) [It is a new masterpiece]

### 2.1.2. Engagement

Engagement in Appraisal implies other voices different from the speaker's, either if they are present or absent in the texts. It deals with sourcing opinions and aligning the speaker, that is, the way in which the author positions him/herself with respect to these opinions as well as hypothetical responses. As Read et al (2007: 94) point out, Engagement provides the resources through which speakers construe their point of view and adopt stances towards others' opinions. These resources include all items "by which the textual or authorial voice is positioned intersubjectively" (White 2001: 14). Speakers have at their disposal a number of options to engage or disengage with opinions, either their own or others', which represent the types of Engagement (e.g. Acknowledge, Counter, Deny, etc.).

Martin and White (2005: 92-94) acknowledge some specific influences in the composition of the axis of Engagement. These are mainly three: Stubbs (1996), Bakhtin (1981) and Voloshinov (1995). Stubbs (1996: 197) assumes that "whenever speakers (or writers) say anything, they encode their point of view towards it", that is, every utterance conveys an opinion. Both Bakhtin and Voloshinov argue that verbal communication takes into account what has been said before, and anticipates what may be said afterwards. In this sense, every utterance is not only opinionated, but also 'dialogic'. However, Martin and White show that there are cases in which other voices are not referred to, such as bare assertions, imperative mood and future tense with *will* and *be going to*. These assertions contrast with dialogistic or heteroglossic utterances in that they do not overtly recognise dialogistic alternatives, but are presented as facts, so they are monoglossic. The distinction between monogloss and heterogloss is the first one for analysing Engagement. All monoglossic expressions are excluded, so the coding scheme is restricted to heterogloss. Figure 3 summarizes all further heteroglossic subsystems.

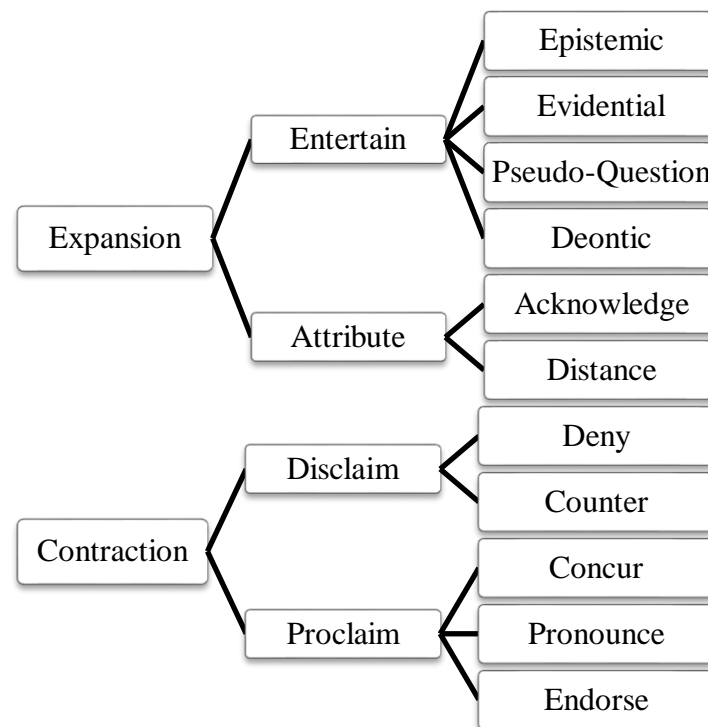


Figure 3. The Engagement system (adapted from Martin and White 2005: 134)

Heteroglossic utterances are subdivided according to

the degree to which an utterance, by dint of one or more of these locutions actively makes allowances for dialogically alternative positions and voices (dialogic expansion) or alternatively acts to challenge, fend off or restrict the scope of such (dialogic contraction). (Martin & White 2005: 102)

Thus, Expansion and Contraction stand for expressions that open or acknowledge different opinions or dialogic voices, and those that limit or challenge those alternative viewpoints.

### 2.1.2.1. Expansion

Expansion offers two ways in which the dialogical voice can be opened up to alternative viewpoints: Entertain resources, which present the author's position as one inside a range of possible options; and Attribute expressions, which provide an external source for a given opinion.

### a) Entertain

Entertain elements are divided into four subtypes: Epistemic, Evidential, Pseudo-Questions and Deontic. Epistemic resources are associated with those traditionally thought to refer to the degree of the authorial voice's certainty about the probabilities and possibilities of something to be true. Martin and White's (2005: 104-5) reinterpretation under the heteroglossic perspective understands these elements as those that recognise the proposition as "one among a number of propositions available in the current communicative context". Examples [33]-[36] show some of these resources.

- [33] Personally, I ... (41)
- [34] It definitely isn't (69)
- [35] No creo que sea necesario (132) [I don't think it's necessary]
- [36] Puede parecer una novela exclusivamente femenina (141) [It may seem an exclusively-feminine novel]

Evidential markers deal with the type of source of evidence the speaker can make use of, as well as asserting that the speaker believes the proposition has (or has not) probabilities to be true, as illustrated by [37]-[38].

- [37] It seemed familiar (50)
- [38] No le veo tanta relevancia (146) [I don't find it so much relevance]

Pseudo-Questions include those resources that do not expect to be answered by the addressee, but they just call into question some information, as [39]-[40] show.

- [39] You want to know why? (77)
- [40] ¿Será que escuchan en vez de leer? (150) [May it be that they listen instead of reading?]

Deontic meanings establish relationships of obligation, control, compliance and permission, traditionally called 'deontic' modality, as can be seen in examples [41]-[42].

[41] The protagonist must confront her demons (78)

[42] Deberías reestructurar el juego (124) [You should restructure the game]

#### b) Attribution

Attribution deals with “those formulations which disassociate the proposition from the text’s internal authorial voice by attributing it to some external source” (Martin and White 2005: 111). Thus, the authorial voice is positioned and expresses their attitude with respect to an external source which is introduced in the discourse. This can be done in two ways: by acknowledging the external source, or distancing from it, typically by means of a reporting item.

Acknowledge is used when writers recognise another source when their point of view is not included in the discourse, that is, they only take into account a different voice, without aligning with it, as exemplified in [43]-[44].

[43] The only thing you would say is... (17)

[44] El que dijo lo de... (126) [The one who said that...]

Distance from others’ propositions appears when they explicitly disagree with them. It must be taken into account that this category may vary according to the meaning found in the content of the whole proposition, which provides a distancing meaning. Examples [45]-[46] illustrate Distance.

[45] Some people say that it is simply. It definitely isn’t. (17)

[46] Pueden decir que es por otra, pero realmente... (148) [They may say it is because of another one, but it really...]

#### 2.1.2.2. Contraction

As said above, dialogic contraction restricts or challenges alternative positions, that is, the authorial voice acknowledges other viewpoints but does not support them. More specifically, it refers to expressions in which speakers disclaim or proclaim a proposition.

### a) Disclaim

Disclaim involves that the authorial voice disagrees or rejects contrary positions from other voices. In turn, it can take the form of Deny and Counter.

Deny is an overt negation of a proposition. Deny also includes items whose meaning, though used in a positive proposition, refers to the non-performance of some action. [47]-[48] illustrate this category.

[47] At the beginning, neither is believable (79)

[48] A mi no me da ningun problema (128) [It doesn't give me any problem]

Similarly to Deny, Counter involves a contrary position to one that has already been introduced. However, it plays with previous expectations about a coming proposition, which is substituted by a different one. Thus, the main meanings that they convey are concession and counter-expectation, as seen in [49]-[50].

[49] However, this is ... (74)

[50] Pero realmente lo unico que quieren...(148) [But what they only really want...]

### b) Proclaim

Proclamation refers to resources used to agree with a proposition from an external source. Thus, the proposition is presented as valid and reliable. Proclamation is achieved through concurrence, pronouncement or endorsement of alternative viewpoints.

Concurring means that there is partial or complete agreement between the authorial voice and the message, or that some knowledge is shared with a majority of voices. For example, [51]-[52] show one of the most common patterns.

[51] Of course she is ... (91)

[52] No voy a negar que oírla es algo maravilloso (186) [I'm not going to deny that listening to her is something wonderful]

Another frequent way of proclaiming others' propositions is Pronounce. When an item can be considered as Pronounce, it means that the author emphasizes or asserts a message. The author may intervene explicitly to express that his or her opinion is firm, without referring to other voices. Examples [53]-[54] illustrate Pronounce.

[53] Believe me I love ... (69)

[54] La verdad es que no es buena (174) [The truth is that it is not good]

Finally, Endorse refers to those propositions from external sources which are presented as correct, unquestionable and undeniable, thus excluding other possible options, as is the case of [55]-[56].

[55] History will show that ... (69)

[56] Como la historia ha demostrado... [as history has demonstrated...]

### **2.1.3. Graduation**

Finally, graduation is the last sub-system used in the analysis. Graduation is concerned with up- and down-scaling, that is, gradability. This system may affect both Attitude and Engagement, inasmuch the former can vary in the degree of positivity or negativity, and the latter can increase or soften the degree of involvement in the utterance (Martin and White 2005: 135-136).

Thus, Graduation deals with two main meanings. On the one hand, it may specify the extent to which an item belongs to a given class, that is, how accurately something can be classified as a member of a semantic category (in other words, prototypicality). On the other hand, it may express the intensity of a feeling. The first of these meanings will be called Focus, while the second will be Force. Figure 4 shows the scheme of graduation, including both Force and Focus and their subclasses.

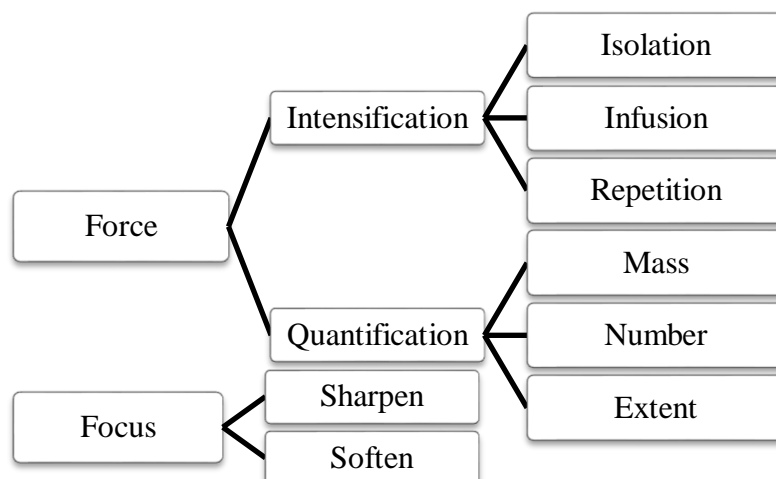


Figure 4. The Graduation system

### 2.1.3.1. Force

Assessments categorised as Force regulate the degrees of amount or intensity. They either control the amount and extent (in proximity and scope) of an entity, typically realised by a noun, or value the intensity of qualities and processes, realised by adjectives, adverbs, verbs and modalities such as probability, usuality and obligation.

#### a) Intensification

Intensification may manifest in three ways. In the first place, a separate word or item can be placed to modulate the intensity of another one. This class is called Isolation and is exemplified in [57]-[60].

- [57] Shame after so long a wait (90)
- [58] That's a pretty huge statement too (98)
- [59] Muy cara (200) [Very expensive]
- [60] El libro es machista a más no poder (147) [The book is sexist to the max]

The second type of Intensification, Infusion, is not realised separately, but it is involved in the very meaning of an item. Its meaning can escalate using a stronger or weaker item, as [61]-[64] show.



- [61] Take a deeper look (69)
- [62] It's always giving me an error editing photos (12)
- [63] Simplemente PERFECTA (167) [Simply PERFECT]
- [64] La manera en que el autor narra la historia es buenísima (160) [The way in which the author tells the story is very-good]

The third and last means to convey intensification is Repetition: this implies that the item to be intensified, or part of it, is repeated. These two possibilities can be seen in [65]-[67].

- [65] What can I say??? (82)
- [66] Way way way too many advertisements!!! (31)
- [67] Me encantaaaa !!! (171) [I loooove it !!!]

#### b) Quantification

When quantifying entities, two aspects can be taken into account: Amount and Extent. Amount refers to factors such as the number, weight, size or strength of the entity. However, Extent refers to different dimensions: space and time. When quantifying the space or time of entities, Quantification statements address how near or far the item is, how recent or old it is, how long it lasts or how widely it is distributed.

Graduation elements about Mass, Number and Extent share the quality of being imprecise. In the case of Mass, examples like those in [68]-[69] can be found. For Number instances, one may look at [70]-[71]. Finally, Extent cases are illustrated by [72]-[73].

- [68] What a colossal selfish pain in the tuckus (83)
- [69] Pero hay un pequeño error (107) [But there is a small mistake]
- [70] A good album with some great tracks (92)
- [71] Súper con algún que otro detalle menor (110) [Great with some minor detail or another]
- [72] And will no doubt appeal to a wide audience (38)
- [73] Llegada reciente

#### 2.1.3.2. Focus

Focus assessments address the issue of prototypicality, that is, how close something is to the core or the ideal item of a category. Expressions inside Focus will graduate how blurred or sharpened the boundary of the category is.

### a) Sharpen

Expressions that sharpen or up-scale an item will place it in the central part of the category it belongs to. For example, this can be seen in instances [74]-[77].

- [74] It's a pure masterpiece (71)
- [75] If you were a true fan, you would love and adore her (T37)
- [76] Un héroe genuino [A genuine hero]
- [77] Un amigo de verdad [a true friend]

### b) Soften

On the contrary, expressions that soften or down-scale an item will place it in the periphery of the category as a marginal member, as examples [78]-[81] show.

- [78] Then it kinda drops off (95)
- [79] The story is new.....ish (T47)
- [80] Una especie de héroe [a kind of hero]
- [81] Es algo así como mi novio [He is sort of my boyfriend]

#### 2.1.4. Previous and related research on Appraisal

The studies related to the analysis of opinion and evaluation expressions in reviews and other texts have grown in the last years. Appraisal Theory has been applied to a wide range of types of texts, including school children's narratives (Coffin 1997; Martin 1996), news reports (White 1998), language acquisition (Painter 2003) and literary texts (Love 2006). Nevertheless, most of these studies focused on a particular set of elements, namely

attitudinal expressions, dismissing evaluative units related to Graduation and Engagement (e.g. Macken-Horarik 2003; Page 2003; White 2004, etc.).

Following the theoretical study of evaluative language, implementations of Appraisal Theory have been carried out over different corpora, annotating and extracting automatically those features. Regarding corpora creation and analysis, Wiebe, Wilson and Cardie (2005), and Wilson and Wiebe (2005) offer a comprehensive annotation scheme for ‘private states’ (Quirk et al, 1985), such as opinions, emotions, etc. Somasundaran, Ruppenhofer and Wiebe (2008) present the concept of ‘opinion frame’ or individual opinions linked by a shared target, in order to overcome sentence-level analysis. With respect to the automatic extraction of these evaluative expressions — also called Sentiment Analysis or Opinion Mining, skilfully reviewed by Pang and Lee (2008) — encouraging results have been obtained. As starting point in computational experimentation with the Appraisal framework, Taboada and Grieve (2004) assigned adjectives into one of the three broad Attitude classes, and presented a prominence schema which revealed that most adjectives are found not in the beginning, but in the middle and the end of online film reviews. Adjectives were also the element of study for Dong (2006) and Wang and Dong (2007), who analysed the semantic orientation of this word class and developed a computational model to identify semantic and grammatical structures of appraisal in texts. Some progresses were also made on automatic extraction of Appraisal in the case of Bloom, Garg and Argamon (2007) and Bloom, Stein and Argamon (2007), again focusing on adjectives as realization of the Attitude axis. More importantly, appraisal taxonomies were shown to improve significantly sentiment extraction and classification in Whitelaw, Garg and Argamon (2005). More recently, Attitude has been discussed as a powerful tool for analysing specialised texts, like in Hommerberg and Don (2015).

There are also specific studies on the linguistic analysis of Engagement. They pay special attention to academic discourse and the acknowledgement of others’ opinion by means of reporting verbs (Chatterjee 2007), although sometimes, despite the similarities, they do not adhere explicitly to the system of Appraisal (Hyland, 1999). Engagement has also shown to be an enlightening tool in the analysis of journalistic and political discourse (White 1998, 2003; Marín-Arrese and Núñez 2006; Miller 2004) and Multimodality (Tan 2010). The work of Brooke et al (2009), who adapted an existing English semantic orientation calculator to Spanish, deserves special mention. Finally, some problems have

also been pointed out regarding overlapping issues between Engagement and Attitude classifications (Carretero and Taboada 2015).

Although Graduation is closely related to the other two axes, it has received little attention in the literature. Only a few studies focus on Graduation as their main research line. Some of the most significant ones are the following ones. In the first place, Vian Jr. (2008) analyses only Graduation realizations in Brazilian Portuguese, presenting some additional important points such as the inclusion of morphological graduation that does not exist in English. Secondly, Carretero and Taboada (2014) carry out a study on examples of Graduation within Attitude comparing English and Spanish reviews. They found out that, although the distribution of occurrences of Graduation is similar in both languages, the use of Affect presents significant differences. Finally, Carretero, Taboada and Hinnel (2014) analyse Graduation and Attitude in three languages (Spanish, English and German) and in film reviews, and point out the high level of Graduation elements found in the texts, highlighting the opinionated nature of the texts.

As a conclusion, although some initial efforts have been made in the analysis of Appraisal Theory beyond Attitude, further study is needed to get new findings on the use, realisations and boundaries of evaluative stances. Moreover, few studies have been carried out contrasting English and Spanish within the Appraisal framework, so a specific emphasis should be made on this aspect, either in monolingual or contrastive analyses.

### **2.2. The genre of ‘Mobile application reviews’**

This study focuses on the genre of 'reviews', and more specifically on reviews of mobile applications. Application here does not only stand for the pure meaning of mobile applications, such as WhatsApp or Instagram, but for any item that can be downloaded or bought via an application store: this includes games, music, books and films as well. Before getting into and outlining the specific characteristics of this growing genre, a brief history of the emergence of application platforms and reviews is presented, followed by a general outline of the notion of genre.

To start with, reviews can be considered subjective texts, inasmuch as they typically present an opinion about the item reviewed. Subjectivity in language refers to those aspects of a language that are used to convey an opinion, speculation or evaluation (Banfield 1982; Wiebe 1994). A very simple definition of sentence objectivity and

subjectivity is stated by Liu (2012: 19): “an objective sentence presents some factual information about the world, while a subjective sentence expresses some personal feelings, views, or beliefs.” He further distinguishes *subjectivity* from the meaning of *opinionated*, which refers to the fact that a sentence or whole document implies a positive or negative sentiment.

This difference is highly relevant for the genre of ‘reviews’, since not only subjective words or expressions convey opinions, but also apparently factual elements may also imply a positive or negative intention. For example, a completely objective statement like “It does not work”, when referring to a mobile application, entails a very negative meaning.

Subjectivity has been recently studied within the area of Sentiment Analysis and Natural Language Processing (NLP), in general. Applications of subjectivity recognition in the NLP community include question-answering (Carbonell 1979), e-mail classification (Aone, Ramos-Santacruz and Niehaus 2000), review mining (Terveen et al. 1997), review classification (Turney 2002; Pang, Lee and Vaithyanathan 2002) and document clustering (Sack 1995).

In the last few years, society has witnessed and starred the emergence of and expansion of the mobile technologies at a blistering pace. Apple Inc. created App Store in 2008, the trailblazing store. Then, Google’s Play Store and Windows Store were launched four years later, in 2012.

The success of app stores is also related to the mass spread of mobile devices. For example, Google recently counted the active Android devices worldwide up to 1.4 billion (Callaham 2015) and Google and Apple’s combined app stores were projected to be worth 25 billion USD in 2015 (marketsandmarkets 2010). Although smartphones were already in use before the creation of app stores, it was in 2008 when they started exploiting their possibilities, which was much fostered by downloadable applications, thus resulting in a symbiotic relationship.

Before application stores, users could only use commercially available applications pre-installed in their mobile devices. However, app stores allow autonomous developers and software engineers to publish their applications and show them to the wider public, who can download it for free or pay for them if established so by the developer.

Therefore, these stores provide a distribution channel for developers to promote their apps as well as a direct feedback channel between users and developers or users and other current or potential users. This can be done by means of a review, a star rating or both. These reviews can either discourage or motivate other users to buy or download the app in question, so they may play a key role in leading the app to the top ranks and make it more visible or fall into oblivion.

In the present study, reviews are presented as a specific and distinctive genre. In order to understand the motivation that leads to be classified as such, a brief summary of the concept of genre is offered below.

Thirty decades ago, long before the creation of app stores, Mikhail Bahktin (1986: 60) referred to the notion of genre in the following way:

Each separate utterance is individual, of course, but each sphere in which language is used develops its own relatively stable types of these utterances. These we may call speech genres.

Bahktin sees language in communication as an individual and concrete realization produced by participants' utterances. These utterances are grouped together, creating types of speech genres. Beyond this brief and general definition, three different perspectives have traditionally addressed genres, focusing on some of their specific aspects and elements: language teaching, especially in English as a Second Language (ESL) (Swales 1981, 1986, 1990), the North American New Rhetoric Studies (Miller 1994) and Systemic Functional Linguistics (Eggins 2004).

The ESL perspective focuses on formal properties and the communicative purposes of genres so that learners of a specific language can master them. The New Rhetoric Studies perspective pays more attention to functional and contextual aspects of genres. Finally, the SFL perspective draws on Halliday's notions of field, tenor and mode to analyse textual features, including form, function and social context. However, these three approaches share some ideas, which are listed by Bhatia (2004: 23). The most relevant aspects of this list for the present study are quoted below:

1. Genres are recognizable communicative events, characterized by a set of communicative purposes [...].

2. Genres are highly structures and conventionalized constructs [...] in terms of the intentions [...], the shape they often take [...] and the lexico-grammatical resources one can employ.
3. Genres [...] focus on social actions embedded within disciplinary, professional and other institutional practices.
4. All disciplinary and professional genres have integrity of their own, which is often identified with reference to a combination of textual, discourse and contextual factors.

After reviewing the three approaches and based on the understanding of them all, Bhatia (2004: 23) presents a comprehensive definition of genre.

Genre essentially refers to language use in a conventionalised communicative setting in order to give expression to a specific set of communicative goals of a disciplinary or social institution, which give rise to stable structural forms by imposing constraints on the use of lexico-grammatical as well as discoursal resources.

Taking this as working definition, some common characteristics can be identified in Google Play Store's reviews that support their conformity as a genre or a subtype inside online reviews.

### **2.2.1. Characteristics of reviews in application stores as a genre**

#### ***Conventionalised communicative setting***

Play Store can be accessed either by their mobile application or the website <https://play.google.com/>. This environment displays the apps users can download together with a list of comments other users have published. These comments are available for any item in any category the app includes, although the less famous an app is, the fewer comments it will have, if at all. All comments are public to either registered or unregistered users, and they can be anonymously written by "a Google user" or no name at all. Figure 5 shows a standard comments display on the PC version.

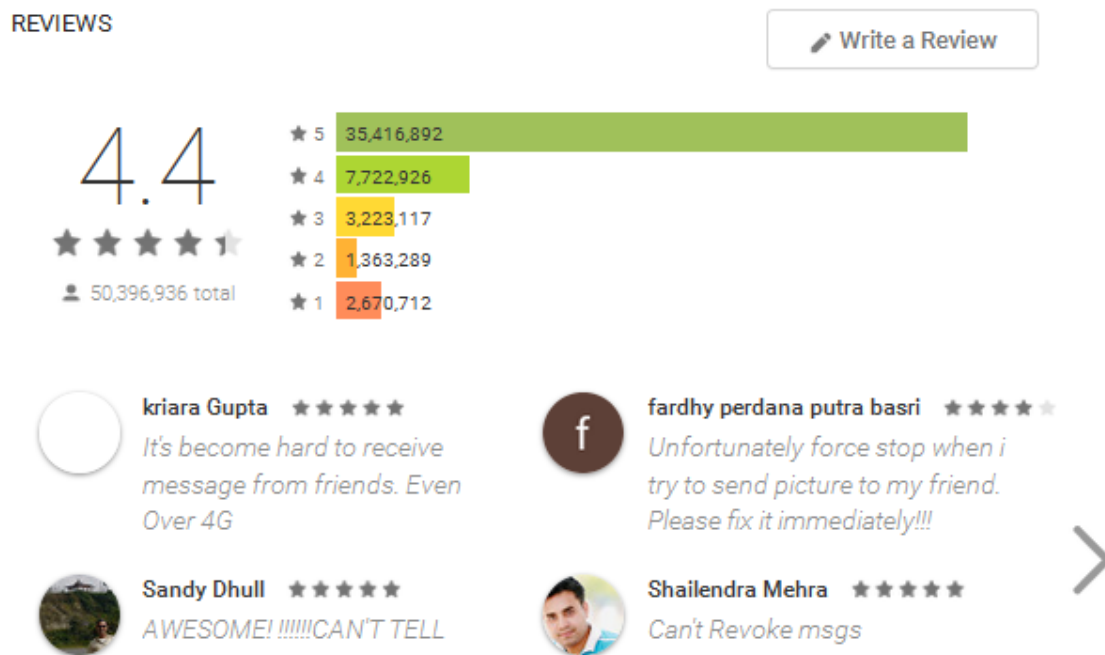



Figure 5. Standard review display (WhatsApp, PC version)

### *Specific set of communicative goals*

When selecting “Write a Review”, the platform provides a guiding text to help users write their comments, as can be seen in Figure 6.

  
WhatsApp  
Messenger  
WhatsApp Inc.

Tell others what you think about this app. Would you recommend it, and why?

★★★★★

CANCEL SUBMIT

Figure 6. Standard review writing display



The main purpose, as suggested by Play Store, is telling other users your opinion about that particular app, and developing further into reasons why you think others should get it or dismiss it.

This may not be different from the original purpose of any other review. However, as can be seen in the very previous figure, users have slightly deviated from valuing the items (as good, bad and other inherently opinionated classifications) and have turned to describing the performance of the items. Three out of the four first comments displayed in WhatsApp profile (Figure 5) address performance issues: messages are not received despite having a good internet connection, it closes when trying to send an image and messages cannot be cancelled.

Another special characteristic that this type of reviews shows is that users address directly application' developers. This can be seen in one of the comments in the figure, where it says "Please fix it immediately!!!" This means that the participants go beyond other users, and include the designers of the item reviewed. When speaking to developers, users may make suggestions for improving the app, requests for fixing bugs or problems or even "blackmailing" them promising a higher rating if the requirements are satisfied (Kimura 2014).

Finally, it is important to mention the apparent randomness of the correspondence between the comments and their star ratings, which has also been noticed in the literature (Islam 2014). Taking another quick look at the comments in Figure 5, one can easily see how three of them post on different problems (no positive aspects described in them) but they rate the app with 4 and 5 stars. However, comments themselves can be rated by other users, as shown in Figure 7:

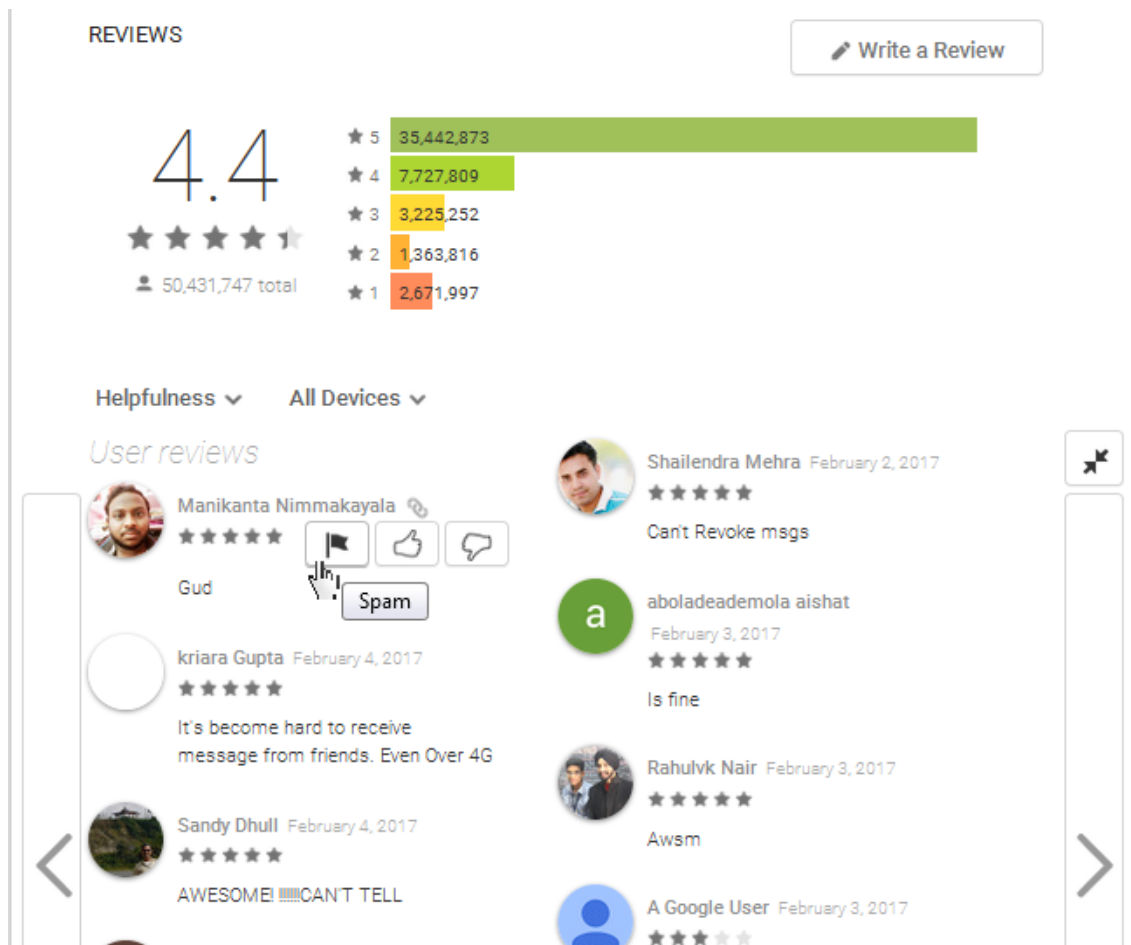


Figure 7. Comments rating by other users

Users can rate another user's comment as "spam", "helpful" or "unhelpful", respectively. This might help solve the problem of ambiguity, but users do not usually spend their time on rating comments.

### *Stable structural forms (constraints on lexico-grammatical and discursal resources)*

The comments are limited to 1200 characters and a 1-to-5 star rating must be selected before submitting. When hovering over a possible star rating, a short clarifying message is shown next to the stars, explaining their meaning:

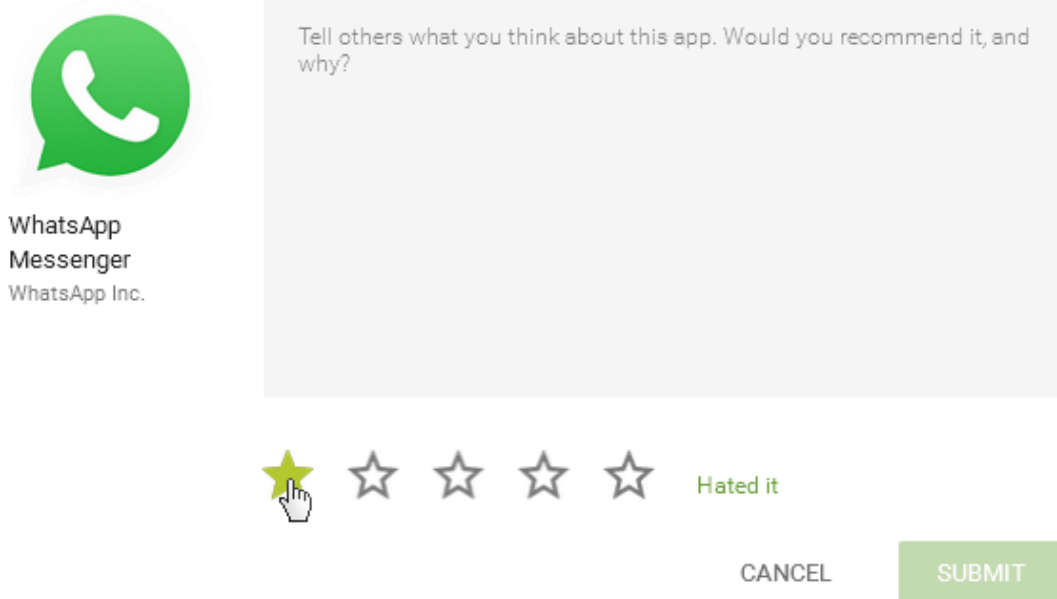


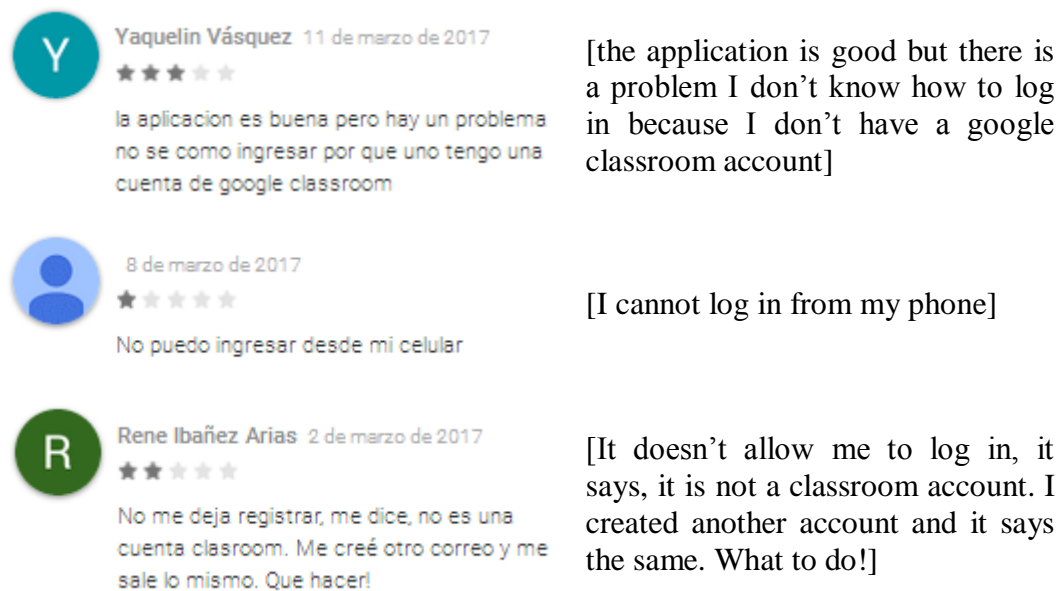
Figure 8. Star rating meaning

The correspondence between the stars and the meaning is the following one: one star, “Hated it”; two stars “Didn’t like it”; three stars, “Just OK”; four stars, “Like it”; and five stars, “Loved it”.

Since comments can be published via smartphone (and, actually, that is the most usual means, instead of using a more traditional keyboard with a PC), typical elements of the internet and mobile language are included, such as abbreviations and emoticons. In addition, sentences frequently miss subjects and links, since authors try to speed up their writing in their phone’s small keyboard.

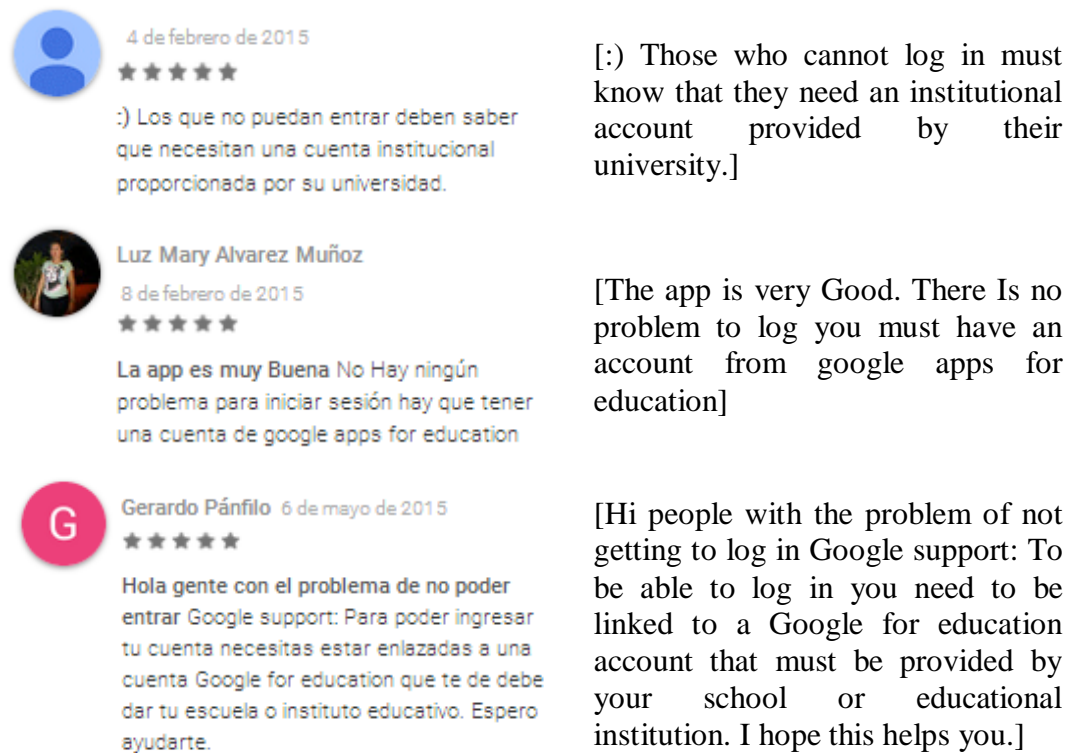
Generally speaking, users do not always follow grammar rules when writing their reviews. A mixture of formal, informal and slang expressions can be found, what makes it even more difficult to find predefined patterns (Khalid, Asif and Shehzaib 2015). The usual profile of a reviewer in application stores is a teenager or young adult, aged 15-35. However, they do not need to be deeply familiarized with new technologies, since some of the negative comments may be due to ignorance in making it work properly. For example, there is an application called Google Classroom, a free service for schools, which is used to create virtual classes, distribute assignments, send announcements and start class discussions, among many other functions. This application requires a specific

Google Classroom account, not the standard Gmail one. However, not every user knows that or how to get that account, so problems are found at the registration step (Figure 9).



**Figure 9. Low rating due to low technological expertise**

As can be seen, users rate the application with 1-3 stars, a negative and neutral rating, and they all blame the problems they are experiencing when introducing their (invalid) account. Actually, other users try to help these ones, since they see the problem does not lie in the application itself.



**Figure 10. Users help other users**

These users have given a 5 star rating to the application because they like the application and they know there are no registration issues.

These are only some of the characteristics that start to delimit the boundaries of this review subgenre. Further conclusions will be drawn after the analysis in chapter 4, once the language of the reviews is analysed.

### **2.2.2. Previous research on reviews published in application stores**

The analysis of opinion and sentiment in the online reviews genre has been usually addressed from the point of view of books (Sorensen and Rasmussen 2004; Chevalier and Mayzlin 2006), films (Chaovalit and Zhou 2005; Kennedy and Inkpen 2006; Zhuang, Jing, Zhy and Zhang 2006; Taboada, 2011; de Jong and Burgers 2013) and hotels (O'Mahony and Smyth 2009; Carrillo de Albornoz, Plaza, Gervás and Díaz 2011) reviews. However, as new technologies are further developed, new types of reviews appear. This is the case of Google Inc.'s Play Store supported by Android devices and Apple Inc.'s App Store, supported by Apple's (iOS) devices. Play Store and App Store are similar applications with similar characteristics, but created upon different operating

systems; thus, the present study deals only with one of them, namely Play Store, which is supported by a wider range of mobile devices.

As has been explained, Play Store is a new platform that allows buying or downloading for free different items (games, applications, books, music, etc.). It has provided the user with a new scenario to share their opinions. The relevant literature on the topic has only been written recently and there is not a significant background on the analysis and measurement of the changes it may be triggering in reviews writing, even though it is more and more used every day. Among the main research done on the topic, the following studies are worth-mentioning.

In Vasa et al (2012), the language of 8.7 million reviews from 17,330 apps was analysed, and they concluded that the comments tend to be short but informative. Moreover, they found out a relationship between the length of the review and the rating: the poorer the rating, the longer the comment.

Khalid (2013) analysed 6,390 reviews for 20 iOS apps in order to extract the type of complaints users made. These complaints included topics such as functional errors, requests for additional features and app crashes.

Guzman and Maalej (2014) analysed 32210 reviews for seven iOS and Android apps and proposed an automated approach to help application developers filter, aggregate and analyse user reviews. Their approach extracted features automatically from the comments that allow developers either to analyse user opinions regarding those single features or to filter irrelevant reviews.

Islam (2014) focused on the ambiguity between opinions content and rating and developed a unified rating system that extracted a numeric polarity from both the text of the review and the star rating and combined them to generate a final rating.

As can be seen, the approach to analysing app reviews is mostly computationally-oriented, based on statistical techniques and quantitative analysis, using word-correlation algorithms and machine learning. However, there are still some pending aspects that have not been deeply studied yet:

- 1) Research on other languages different from English is scarce (if at all), so contrastive studies have not been exploited yet.
- 2) Research on linguistic features (beyond the topic or features mentioned in the review) does only cover superficial filtering such as lemmatisation or content words (noun, verbs and adjectives, mainly) extraction.

- 3) Previous research has focused on applications, leaving aside comments on other items such as books, music and films on the same platform, which may present different characteristics either from application reviews or more traditional reviews on other platforms (e.g. epinions.com and the like).

Therefore, the present study provides a linguistic analysis that can help other approaches improve their performances or move forward in the understanding of this genre.





## Chapter 3

# VALIDATING APPRAISAL FEATURES THROUGH CORPUS ANNOTATION

### **3.1. Introduction**

This chapter addresses one of the main aims of this dissertation, namely, to validate aspects of the Appraisal Theory empirically and cross-linguistically. Following the methodology proposed by Lavid (2012) and Hovy and Lavid (2010), the idea is to use the results of manual corpus annotation as a mechanism to test hypotheses about linguistic phenomena empirically. In fact, many hypotheses about linguistic phenomena remain as such, i.e., untested empirically. Therefore, using corpus annotation to test aspects of Appraisal Theory empirically is a methodological innovation in current linguistic practice. Also, the creation of empirically-validated annotated resources is a much-needed task for many NLP applications, where the training of computer algorithms must be based on high-quality, consistent, and empirically-validated annotations.

In the following sections I will describe the methodological steps which have been taken to use corpus annotation for the validation of a number of aspects of Appraisal

Theory empirically, following the methodology proposed by Hovy and Lavid (2010), and Lavid (2012):

- (1) Selecting the corpus. This step involves the compilation of what is known as the ‘training corpus’, i.e, a small sample from the larger dataset.
- (2) Delimiting the category(es) to be annotated and its/their features.
- (3) Designing the annotation scheme and guidelines. This involves instantiating all or part of the features of the selected theoretical model and developing a core and an extended tagset to be used in the process of annotating the training corpus.
- (4) Performing annotation experiments on the training corpus in order to determine the feasibility and consistency both of the instantiation and the annotator manual.
- (5) Measuring the results of the annotations by comparing the degree of agreement between the coders' decisions. This step also involves deciding which measures are appropriate, how they should be applied, and determining the level of agreement which will be considered satisfactory.

Each of these steps is described in detail in the subsections below.

## 3.2. The ‘training corpus’

The training corpus is a small subset of the total number of texts extracted from the well-known mobile application and website Google Play. In this app or website, people cannot only download applications for their mobile devices but also write an opinion or comment about the downloaded item in question, as was explained in section 2.2.1.

The training corpus consists of 50 texts that were selected from the total amount of texts extracted in the first step of the process. In order to compile the corpus, four characteristics were taken into account:

1. Language: English or Spanish.
2. Item or field: application, game, book, music, film.
3. Polarity: positive or negative.
4. Rating: five, four, two or one stars.

Tables 1 and 2 show a plain distribution, language, polarity and rating of the 50 selected texts (*st.* stands for ‘stars’, regarding the 1-to-5 star rating).

Applications				Games				Books				Music				Films			
5				5				5				5				5			
+		-		+		-		+		-		+		-		+		-	
2		3		3		2		2		3		3		2		2		3	
5st.	1	1st.	2	5st.	2	1st.	1	5st.	1	1st.	2	5st.	2	1st.	1	5st.	1	1st.	2
4st.	1	2st.	1	4st.	1	2st.	1	4st.	1	2st.	1	4st.	1	2st.	1	4st.	1	2st.	1

Table 1. Selection of 25 English texts for the training corpus

Applications				Games				Books				Music				Films			
5				5				5				5				5			
+		-		+		-		+		-		+		-		+		-	
2		3		3		2		2		3		3		2		2		3	
5st.	1	1st.	2	5st.	2	1st.	1	5st.	1	1st.	2	5st.	2	1st.	1	5st.	1	1st.	2
4st.	1	2st.	1	4st.	1	2st.	1	4st.	1	2st.	1	4st.	1	2st.	1	4st.	1	2st.	1

Table 2. Selection of 25 Spanish texts for the training corpus

As can be seen, the corpus contemplates all the categories and number of comments distributed evenly in order to provide a representative sample. When an odd number of texts had to be selected, the preferred ratings were the highest (5 stars) or the lowest (1 star) ones, since they are more likely to include a higher number of Appraisal realizations, being their opinions more polarized.

The comments in the training corpus address a single item per category, due to the small number of each of them. These items are the following ones:

1. WhatsApp (application): an instant messaging application.
2. Candy Crush (game): a match-three puzzle game.
3. *Fifty Shades of Grey* (book): an erotic romantic novel by E. L. James.
4. Adele (music): a soul, pop and R&B singer.
5. *Frozen* (film): a 3D computer-animated musical fantasy film produced by Disney.

The selection of the corpus was the result of a process of extraction and filtering. In the first place, thousands of comments were extracted automatically with a crawler designed *ad hoc*. This crawler was programmed specifically for that purpose: it performs HTTP GET petitions, that is, it acts as a user checking each comment individually, but distributed through time to keep the queries below the limits established by Google. The crawler obtains data in JSON format in these HTTP GET petitions, parses the JSON

obtained, extracts the comments and metadata in that JSON, and stores everything in a SQLite database for later analysis, like dumping the comments to CSV files.

The amount of comments extracted is displayed in Appendix C, but a summary is presented in Table 3.

	<i>English</i>	<i>Spanish</i>
<b>Applications</b>	15721	15225
<b>Games</b>	15288	15328
<b>Books</b>	4909	2223
<b>Films</b>	7793	1595
<b>Music</b>	5976	2933
<b><i>Total</i></b>	<b><i>49687</i></b>	<b><i>37304</i></b>

Table 3. Summary of English and Spanish comments extracted

### 3.3. Appraisal categories selected for annotation

The Appraisal categories selected for the annotation of the training corpus were the three main divisions: Attitude, Engagement and Graduation, as graphically displayed in Figure 11.

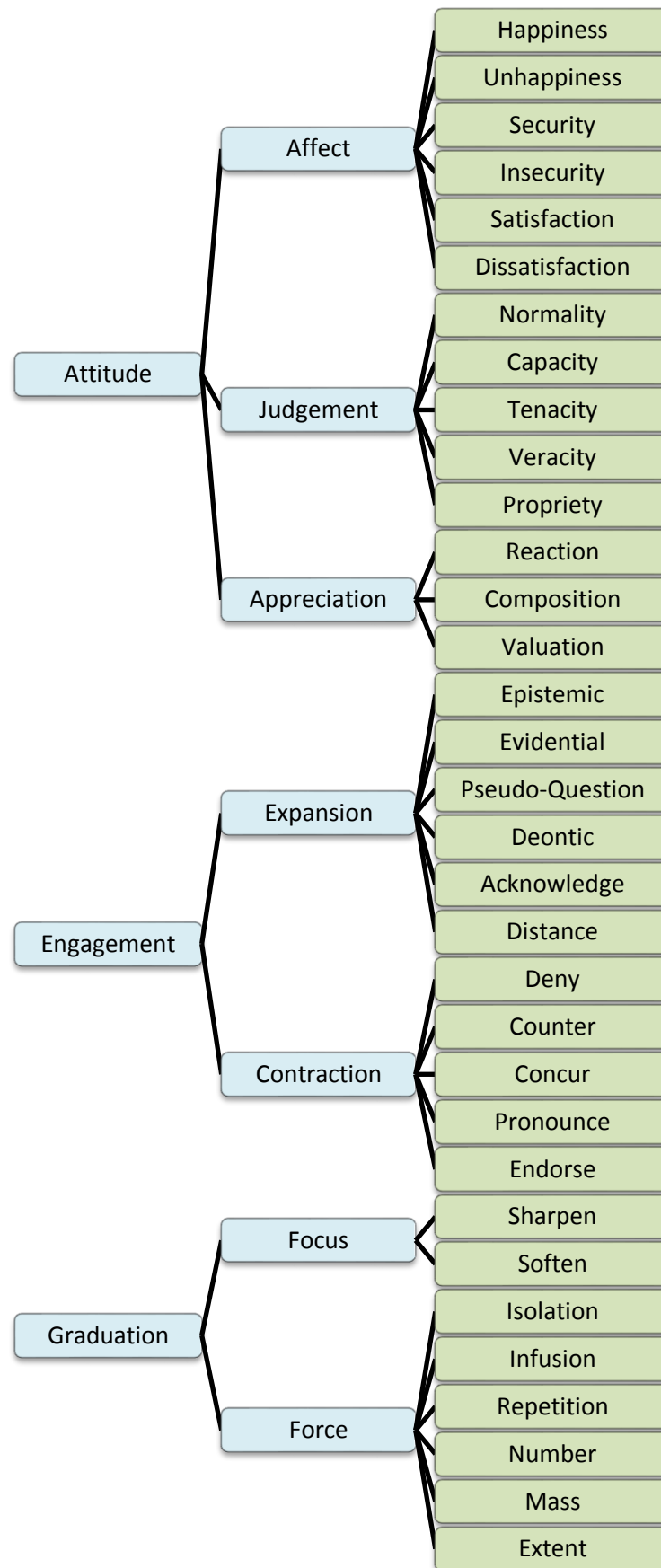


Figure 11. Appraisal categories selected for annotation

These categories are described in the subsections below.

### 3.3.1. Attitude

Attitude deals mainly with feelings, including emotional reactions, judgements of behaviour and evaluation of qualities of things. These three types of meanings represent the three main subtypes of Attitude, namely Affect, Judgement and Appreciation.

Affect includes feelings in their most pure sense: feelings of Happiness, Unhappiness, Security, Insecurity, Satisfaction and Dissatisfaction. Although these feelings are usually realised by adjectives (*sad, glad, afraid, disappointed*, etc.), they can also be found in verbs (*miss, smile, laugh, like*, etc.), nouns (*anger, love, surprise*, etc.), or adverbs (*sadly, restlessly*, etc.).

Judgement addresses meanings of aspects that are socially praised or socially sanctioned. It does not exactly adhere to feelings as such, but they are more quality-oriented. These qualities stand for Normality (how usual or unusual someone is), Capacity (how capable or incapable) and Tenacity (how resolute or irresolute), on the one hand, and they address Veracity (how truthful or disloyal) and Propriety (how moral or immoral). Like feelings in Affect, these meanings are usually applied over people. However, they can also be used metaphorically when, for example, an ATM does not give us the money it is charging us, and we say it is *greedy*, or if it gives us more money by mistake instead, we can say it is *generous*. Therefore, Judgement includes meanings that assess behaviour, either people's or things'.

Appreciation provides qualities for the value of some addressee. This is the category that is typically used when talking about something else, and can be used in three ways: Reaction, Composition or Valuation. Reaction gives the object reviewed the power of causing something, usually a feeling. Thus, the difference with Affect is that the focus is not placed on the feeling produced and who suffers it, but on the item that causes it. When this comes to language, it is typically (but not only) realised by adjectives ending in *-ing* in English, and *-ante* or *-(ad)or* in Spanish, such as *amazing* (it amazes you), *fascinating* (it fascinates you), *espeluznante* [horrifying] (it horrifies you) or *perturbador* [disturbing] (it disturbs you). Composition evaluates the complexity and balance of the item: something may be *difficult* or *easy, simple, unclear*, etc. Finally, Valuation stands for mostly any other quality implying an evaluation, that is, a *blue-eyed*

girl cannot be considered as evaluative, but an *original* girl can. Concepts like *good* or *bad* (when not referring to moral ideas), have also been included in this category, since they cannot be said to express a Reaction nor a Composition but, despite the infinite ways in which they can be used, they typically belong to a more general Valuation on the item.

### 3.3.2. Engagement

Engagement attempts to reveal the authorial voice as one of a range of possible voices and alternative viewpoints that are included (either explicitly or implicitly) in the text. Statements like imperative sentences (*Take it*) do not allow any other voices, since they are factual expressions. However, there are a substantial number of resources that involve Engagement issues, some of them present the author's statement as just one possible one related to individual subjectivity (Expansion), while others restrict the scope of these other alternative positions and voices (Contraction).

In the first group, six categories are analysed. The first one, Epistemic, basically addresses the level of certainty that authors provide their statements with. For example, *I think* or *I'm sure* represent Epistemic realisations with different degrees of certainty, and the same happens with *They may come* or *They might come*. Secondly, Evidential markers introduce an empirical source for the information communicated. The usual source is the sight, so typical examples include *It seems* or *it appears*, but also *it sounds*. Thirdly, Pseudo-Questions cover all those questions which are not intended to be answered by the interlocutor, so they are usually employed to express rejection or surprise on some facts. Fourthly, Deontic resources convey obligation, control, and resistance meanings. These are usually realised by *must* (when it does not imply deduction) and *have to*. Fifthly, Acknowledge introduces an external source of opinion but authors do not position themselves either in agreement or disagreement with it. And lastly, Distance does also introduce an external source, but the author shows disagreement. It must be taken into account that these two last categories may not have a specific lexical item to introduce them (any reporting verb would be probably valid), but the author's position can be found in contextual information.

The second major category, Contraction, is divided into five additional subtypes. Deny is a common resource in language in general. It involves all elements implying negation, not only verbal negation such as *not*, but also others such as *no*, *none*, *nothing*,

*never*, etc. Counter expresses concession or counter expectation, usually in the form of conjunctions and connectives (*although*, *but*) and comment adjuncts and adverbials like *surprisingly*, *only*, *even*, etc. Concur and Endorse introduce agreement with external sources but, while Concur only indicates there is such an agreement (*of course*, *naturally*, *obviously*, etc.), Endorse also presents the statement as undeniable (*Johnson has demonstrated that...*).

#### 3.3.3. Graduation

The third axis, Graduation, provides resources to modulate the meanings introduced by Attitude and Engagement elements. This modulation is done either by Force or Focus items. Force deals with the degree of intensity or amount of the item addressed. This category is subdivided into some others. The first one, Isolation, is the most basic and common one, including lexical items that perform that scaling individually and are attached to the modulated item, as words like *very* or *so* do or punctuation marks like exclamation ones. Infusion, in contrast, involves terms whose scaled meaning is inserted in the very same term. This is the case of comparative adjectives like *better* (in opposition to *good*, which has a lower intensity) or capital letters, which manifest their intensity in opposition to small letters. Repetition is another resource, and can be used either by repeating a full lexical item (or a combination of items), as in *very very very rich*, or part of it, as in *I loooooove it*. Number includes all those indefinite numerical modifiers that graduate the quantity of an item. For example, it is not the same that an application has *many mistakes* as *some mistakes* or even *few mistakes*. Mass addresses the size of an item, like in *it is a big mistake* in contrast with *it is a small mistake*. Finally, Extent covers meanings about proximity and distribution in time and space, like in *a recent update* or *a wide-spreading fashion*.

On the other hand, Focus addresses the item's degree of prototypicality. It can be done either by Soften or Sharpen resources, with opposite results. Sharpen maximises the degree of prototypicality by presenting the modulated as a core one, one that perfectly covers the qualities and characteristics of the meaning it stands for (*he is a genuine hero*), while Soften presents that meaning as a marginal one inside its group (*he was kind of depressed*).



### 3.4. Annotation scheme and guidelines

As has been defined in the literature, an annotation scheme is “the document describing and explaining the scheme of analysis employed for the annotations” (Leech 1997: 6). The process underlying these descriptions and explanations requires a fine-grained simplification of the very abstract concepts presented in Appraisal Theory.

The annotation scheme was divided into a core tagset, consisting of coarser tags, and an extended tagset, consisting of more delicate ones. Both tagsets are graphically displayed in Tables 4 and 5.

Attitude	Feelings, including emotional reactions, judgements of behaviour and evaluation of qualities of things.
Affect	Emotional reactions and feelings
Judgement	Assessment of behaviour according to normative principles
Appreciation	Evaluation and valuation of things
Engagement	Implication of other possibilities and voices than the speakers’
Expansion	Author’s position is one inside a range of possible options or an external source provided for a given opinion
Contraction	Author positions against a contrary position or limits the scope of possibilities
Graduation	Grading phenomena whereby feelings are amplified or softened and categories blurred or sharpened.
Focus	Degree of prototypicality
Force	Degree of intensity or amount

**Table 4. Core tagset for Appraisal analysis**

Happiness	Feelings of cheer and affection
Unhappiness	Feelings of misery and antipathy
Security	Feelings of confidence and trust
Insecurity	Feelings of disquiet and surprise
Satisfaction	Feelings of interest and pleasure
Dissatisfaction	Feelings of ennui and displeasure
Normality	How (un)usual someone or something is
Capacity	How (in)capable someone or something is
Tenacity	How (ir)resolute someone or something is
Veracity	How truthful someone or something is
Propriety	How ethical someone or something is
Reaction	Effect, impact and appearance
Composition	Balance and complexity
Valuation	Value and worthiness
Epistemic	Degree of certainty
Evidential	Empirical evidence
Pseudo-Question	Questions with unintended answer
Deontic	Obligation, control and compliance
Acknowledge	External source provided for a given opinion with no overt indication as to where the authorial voice stands
Distance	External source provided for a given opinion with an explicit distancing of the authorial
Deny	Negation
Counter	Concession and counter expectation
Concur	Overt announcement of agreement or shared knowledge with a projected alternative
Pronounce	Authorial emphases
Endorse	External sources construed by the authorial voice as correct and undeniable
Sharpen	Maximisation of the degree of prototypicality
Soften	Lessening of the degree of prototypicality
Isolation	Scaling of intensity realised by an individual lexical item
Infusion	Scaling of intensity realised by one aspect of the meaning of a single term
Repetition	Repetition of the same lexical item or part of it
Number	Numerical quantifiers
Mass	Extension of an item's presence in space
Extent	Proximity and distribution in time and space

**Table 5. Extended tagset for Appraisal annotation**

The specific guidelines that coders were provided with are included in Appendix D.

### **3.5. Annotation experiments**

On the basis of the tagset proposed in the annotation scheme, three experiments (also called ‘agreement studies’) were designed to test the reproducibility of the proposed tags. The first experiment focused on the identification of the spans or markables, the second one addressed the selection of a broad or general category, and the third one asked to select a specific subtype. These are described in detail in the following subsections.

#### **3.5.1. Agreement study 1**

The purpose of this first experiment was to investigate which elements were considered as Appraisal tags by two independent coders and to delimit their boundaries. As a first approach, every single element in the texts was eligible for conveying Appraisal. That means no part of the text was disregarded *a priori*, no matter whether they talked about the protagonist or the argument of a book, for example, instead of the book itself.

This was because the texts were very short and all the information provided adds a value to the comment, and the interpretation of the argument or the characters themselves may actually well be biased by the author of the review or comment. Moreover, the texts address specifically authors’ opinions, they are not written to offer general or neutral information for other users. Only emoticons were disregarded because of the problems associated with text extraction and conversion to different formats other than the original one, as well as the text processing limitations of the tools employed.

Regarding the length of the span, coders were instructed to annotate the shortest lexical span expressing Appraisal. For example, when annotating negation in contracted forms, coders were asked to annotate the whole auxiliary form (e.g. *won’t*) but to annotate only the negative item when the form was analytical (e.g. *will not*). However, when a contraction could be clearly separated from a pronoun (e.g. *you’d*) or verb (e.g. *would’ve*), only the contraction was supposed to be marked (*’d* and *’ve* respectively). Also, coders were told to annotate word combinations which could be considered as different Appraisal expressions: for example, *I think* would be annotated as an Epistemic marker, while *You think* would be given a different tag (Acknowledge).

The length of the spans covered from the first word to the last word relevant for the span. This means that some unrelated words may have been included in the span. For example, a span like *You think* may have been delimited as *You may really think*, in which *may* and *really* have different purposes from the single combination of *you* and *think*. However, there were some constrictions on the selection of the markables as well, as explained below.

Proper nouns, even if including lexical items that would have otherwise been clearly Appraisal targets, were not given a label. For example, *Sweet Calendar* (T17) is the name of a specific calendar inside the game Candy Crush and, consequently, *Sweet* is not working as an evaluative term. In the same way, titles of books, songs, albums, etc. were not annotated, since they do not constitute statements but are merely referential, like proper nouns.

Parts of fixed and idiomatic expressions and greetings were not annotated even if they included items that would have otherwise been labelled. For example, in the case of an expression like *dejo mucho que desear* [(it) left much to be desired] (T23), potential markables like *best* or *mucho* [much] have not been analysed because they are not fulfilling their usual function but just working as a part of the full expression, that is, *dejar mucho que desear* [leave much to be desired] is not used as *dejar que desear* [leave to be desired], so *mucho* [much] is not an Isolation marker introduced by the author to emphasise their comments. However, other items showing gratitude or congratulations such as *thank you* (10) or *felicitaciones* (T31) were annotated, since they show authors' feelings.

Informative or error messages whose source was the application itself were not annotated either. For example, the words *owner may have changed passwords* in [82] were disregarded.

[82] Ig automatically signs me out of my accounts for some reason saying.

"owner may have changed passwords" (8)

Other elements which were disregarded include clearly representative adjectives or representative verbs. This is the case of *new* (T9, T16) or *old(er)* (T9) when referring to updates or versions of the applications, as can be seen in [83], [84] and [85], or *can* when involving a capability meaning (T7, T9), like [86] and [87] show.

- [83] New update is rubbish. (T9)  
[84] I am so happy that you have so many new levels. (T16)  
[85] prefer the older one (T9)  
[86] Love how you can chat from your computer- its super convenient. (T7)  
[87] can never hear the person on the other end (T9)

### 3.5.2. Agreement study 2

Once annotators agreed on the spans to be annotated, the second annotation experiment addressed the labelling of one of the Appraisal markable with one of the three coarse tags, i.e.: Attitude (Affect, Judgement and Appreciation), Engagement (Expansion and Contraction) and Graduation (Force and Focus).

The purpose of this experiment was to investigate whether the annotators could distinguish between the coarse tags, before getting deeper into more delicate categories. If significant inconsistencies were found, this step would make it easier to identify any conflictive and confusing aspects in the theory or the guidelines.

During the annotation process, possible misspellings or casual mistakes, as in [88] and [89], were annotated as their intended meanings. There were also other common misspellings or abbreviations ([90]) but, in other cases, the author's meaning needed some deeper interpretation, as in [91], where the interpreted meaning is "pdf form/format is only supported is another another".

- [88] me lo e vuelto s descargar (T11) [(I) ave downloaded ot again]  
[89] if you were a true fan you would love and adore her NO matter want kind of album she made (T37)  
[90] 2 many (22)  
[91] PDF from is only supposed is another issue (T6)

In these cases, the texts were annotated as their intended meanings, instead of the actually written words.

### 3.5.3. Agreement study 3

In the third annotation experiment, coders were instructed to use a more fine-grained tag from the extended tagset to label the selected markable. These include Happiness, Unhappiness, Security, Insecurity, Satisfaction and Dissatisfaction, in the case of Attitude; Normality, Capacity, Tenacity, Veracity and Propriety, in the case of Judgement; Reaction, Composition or Valuation, in the case of Appreciation; Epistemic, Evidential, Pseudo-Question, Deontic, Acknowledge and Distance in the case of Expansion; Deny, Counter, Concur, Pronounce and Endorse in the case of Contraction; Sharpen and Softer in the case of Focus; and Isolation, Infusion, Repetition, Number, Mass and Extent in the case of Force.

The purpose of this experiment was to investigate whether highly delicate categories are coded consistently by two independent coders, thus being able to identify subtle differences in meaning and intention in the authors' words. Moreover, tags with blurred differences could be detected and possible problems regarding theoretical aspects may come into view.

## 3.6. Evaluating the annotations

The statistical test used to measure inter-annotator agreement in the three studies (span selection, major category selection and minor category selection) was the *kappa coefficient* (K). Kappa measures agreement among coders when at least two independent annotators analyse the same item. The operation compares the actual agreement to the expected agreement by chance and provides a reliability value. This value moves gradually from  $\leq 0$ , when the agreement is no other than the expected by chance, to 1, if there is total agreement. Table 6 shows what the K values stand for.

<u>K-Value</u>	<u>Agreement</u>
< 0	Less than chance agreement
0.01–0.20	Slight agreement
0.21– 0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–0.99	Almost perfect agreement

**Table 6. Interpretation of *Kappa* (from Viera and Garret 2005: 362)**

The k-value is obtained following the formula:

$$k = \frac{\text{Pr}(a) - \text{Pr}(e)}{1 - \text{Pr}(e)}$$

where  $\text{Pr}(a)$  stands for the relative observed agreement among coders, that is, accuracy, and  $\text{Pr}(e)$  stands for the hypothetical probability of agreement by chance. The results obtained after applying the Kappa coefficient to each of the three experiments will show the level of agreement for each analysis, as will be explained in the following sections. In addition, the cases and reasons for disagreement will be discussed as well.

### **3.6.1. Annotation results for agreement study 1**

Agreement study 1 focused on the identification of Appraisal markables. For the sake of simplicity and convenience, the texts were divided into sentences, using the same templates for the three agreement studies. Then, every time a coder found an Appraisal item, ranging either from a single word, to a phrase or even the full sentence, he/she had to mark it as such. If the same item could be annotated more than once, like in the case of comparative and superlative adjectives, which have a value related to their meaning as well as a value because of their comparative and superlative forms, it should be marked twice, as two occurrences. Later, the number of occurrences per sentence for each coder were collected and compared, and these results are displayed in Table 7.

		Annotator 1											
Annotator 2	Occ	0	1	2	3	4	5	6	7	8	9		
	0	24	2	0	0	0	0	0	0	0	0	26	8%
	1	0	74	9	0	0	0	0	0	0	0	83	27%
	2	0	1	78	7	1	0	0	0	0	0	87	28%
	3	0	0	0	54	6	0	0	0	0	0	60	20%
	4	0	0	0	0	29	2	0	0	0	0	31	10%
	5	0	0	0	0	1	9	1	0	0	0	11	4%
	6	0	0	0	0	0	0	3	0	0	0	3	1%
	7	0	0	0	0	0	0	1	3	1	0	5	2%
	8	0	0	0	0	0	0	0	0	0	0	0	0%
	9	0	0	0	0	0	0	0	0	0	1	1	0%
	Total	24	77	87	61	37	11	5	3	1	1	307	100%
		8%	25%	28%	20%	12%	4%	2%	1%	0%	0%	100%	

**Table 7. Inter-annotator agreement in span selection**

The  $Pr(a)$  value is 0.89 and the  $Pr(e)$  is 0.2, obtaining a k-value of 0.86, a substantially high result.

After analysing the statistics obtained with the Kappa coefficient, a list and explanation of the cases in which disagreement was found is presented below. Figure 12 shows the percentage of cases in which coders did not select the same markable, that is, the span was either selected by one of the annotators and not by the other one, or the span's length was different, in comparison with the total number of spans annotated.



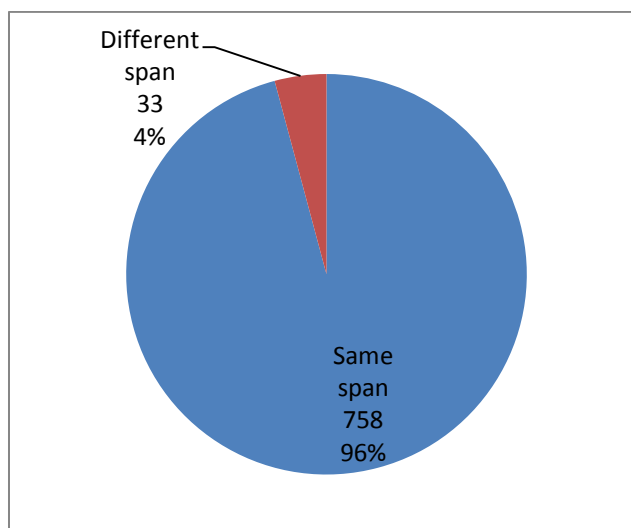


Figure 12. Distribution of agreement and disagreement in markable identification

In most cases the disagreement was found in one annotator selecting a span while the other one did not select it at all. Only once the difference was found in the length of the span. All the spans listed were decided to be annotated except when indicated otherwise.

*Cada vez* [every time] (T2, T5): it either emphasizes the following element (*Cada vez va peor* (T5) [It works worse and worse]) indicating a growing tendency or is used as a synonym for *always* (*Cada vez que la actualizan* (T5) [Every time it is updated]).

*Detalles* [Details] (T3): it assigns a specific value to some issues, that is, it assigns a level of importance to the aspects commented.

*Issue* (T6), *issues* (T19): unlike *detalles* [details], *issue* or *issues* can be understood in two different ways. On the one hand, it can be used as a synonym for *problem* (case in which it would be marked) or, on the other hand, as just referring to topics or aspects of a particular item. Due to the ambiguity of the term in the annotated texts, it was decided to be disregarded, since none of the two meanings was clearly predominant or, in any case, the descriptive instead of evaluative meaning was interpreted.

*Everytime I'm on a whatsapp call always says 'reconnecting'* (T9) and *didn't have this problem with the older version so no its not my phone or network* (T9) y *problem* (T9): they were disregarded because they are merely describing the way the application works and any evaluative meaning is beyond the scope of the textual elements.

*Te pone a trabajar la mente* [it gets your mind working] (T13): it is specifying the effect the game has on the player. It could be clearly opposed by words like *useless* or *boring*.

*Was created by Cain's people* (T17): this expression is assigned a kind of evil value to the game (even though it is used in a positive comment) because it is addictive, etc.

*Película Disney* [Disney film] (T24): initially, the meaning was not clear enough. It can either present a description or an evaluation. This second interpretation was the annotated one because of the connotations that can be found in it, these being innocent, kid-friendly, childish, etc. as a clear opposition to an erotic film, as the author supposed *50 Shades of Grey* (the evaluated film) should be.

*Couldn't put it down* and *in one day* (T26): although these expressions are describing an action performed by the author (the author couldn't stop reading the book and it was read in one single day), they are clearly meaning that the book was engaging.

*Breaking me from the book* and *couldn't put it down* (T27): these spans present exactly the same situation as the previous ones, so they were also annotated for the same reasons.

*In one sitting* (T28): once again, the span refers to the engaging quality of the book commented.

*A 16 years old being brainwashed by some rich man* (T29): although it may seem this span is describing the argument of the book it refers to, it is actually paraphrasing it. This interpretation shows the author's opinion about what is going on in the story and it is evaluated in a negative way.

*I was wondering* (T29): although it is not a prototypical one, it can be considered a way to introduce an indirect question (or pseudo-question) in the text.

*Música para los oídos* [Music to my ears] (T32): the disagreement was caused by this being a kind of fixed expression, but it is assigning a clear value to the reviewed item.

*Vale la fama que tiene* [worths its popularity] (T43): this is the only case in which the disagreement was found in the length of the span. While one annotator selected the full phrase, the other one selected only the verb *vale* [worths]. The full span was finally annotated, since similar cases like *valer la pena* [be worth it] were annotated as a single item.

*Mas que* [more than] (T44): an equivalent meaning for *mas que* [more than] in the comment is *only*. Therefore, it was annotated in the same way as *only* would.

*Bring anything new to the table* (T48): due to the clear identification of this span with an adjective like innovative, it was finally annotated.

*Victoria secret model* (T49): although it refers to an entity, it establishes a comparison between one of the characters of the film reviewed and a model with either particular aesthetical or moral characteristics that are assigned by the author's point of view.

*What the hell is so fascinating about it* (T49): as happened with *I was wondering* (T29), this span introduces a kind of unanswered or pseudo-question.

*What the hell* (T49): as an exclamatory expression, it provides the sentence with emphasis.

*Destroy my TV, tear my ears off and gauge out my eyeballs* (T50): it talks about the feeling the film produces on the reviewer. Instead of assigning a word for that feeling (for example, wrath), the author specifies the actions that could be used as a consequence of that feeling.

As can be seen, most of the disagreement was found in long and complex sentences that do not directly reflect an opinion, but must be contextualised to convey an evaluative meaning. There is a very thin line separating objective descriptions or representative elements from evaluative ones assessing other entities. Moreover, these usually long spans cannot always be easily translated into a single item or term, which makes it more difficult for annotators to agree on their connotations. Another important aspect to be taken into account is the entity they refer to and its qualities. This is clearly exemplified by *película Disney* [Disney film] (T24), where the key to annotate it as an evaluative element is the reviewed film.

### **3.6.2. Annotation results for agreement study 2**

After the first agreement study, a final selection of the markables was selected, as has been explained in the disagreement cases. Once these cases were clear, the second agreement study was carried out. In this case, annotators had to select a major Appraisal category (Attitude, Engagement or Graduation) and a major subtype (Affect, Judgement, Appreciation, Expansion, Contraction, Force or Focus).

The collection of results followed the same procedure as the first one, numbering the occurrences of each value in the spans previously selected. The results are displayed in Table 8.

		Annotator 1										
		Aff	Jud	App	Ent	Att	Dis	Pro	For			Foc
Annotator 2	Aff	82	0	6	0	0	0	0	0	0	88	11%
	Jud	0	45	2	0	0	0	0	0	0	47	6%
	App	1	6	189	0	0	0	0	1	0	197	25%
	Ent	0	0	0	32	0	0	0	0	0	32	4%
	Att	0	0	0	0	9	0	0	0	0	9	1%
	Dis	0	0	0	0	0	178	0	0	1	179	23%
	Pro	0	0	0	1	0	0	13	1	0	15	2%
	For	0	0	0	0	0	0	0	201	0	202	26%
	Foc	0	0	0	0	0	0	0	4	4	8	1%
Total	83	51	197	33	9	178	14	207	5	777	100%	
		11%	7%	25%	4%	1%	23%	2%	27%	1%	100%	

*Just* (T10): it was annotated as Engagement-Contraction and Graduation-Focus. The Focus interpretation was the agreed one because it is delimiting the action, that is, it refers to accuracy instead of being a synonym of *merely*.

*That* (T10): it was annotated as Graduation-Force and Graduation-Focus. It is intensifying the adjective *stupid* and could be replaced by *so*. Therefore, it was finally annotated as Force.

*Recomiendo* [(I) recommend] (T11, T12, T36): it was annotated as Attitude-Appreciation and Attitude-Affect. It expresses a feeling of satisfaction with the product that is manifested in the form of a recommendation. It is the author who feels this satisfaction and performs the action of recommending, so it was consequently annotated as Attitude-Affect.

*Así de* [that] (T11): it was annotated as Graduation-Force and Graduation-Focus. It is the Spanish equivalent for the case of *that* functioning as *so* explained above. It was similarly resolved as Force.

*Shakes the old cobwebs off your brain* (T18): it was annotated as Attitude-Affect and Attitude-Appreciation. It refers to something that the game causes inside the player after playing the game, that is, a quality of the game, so it was agreed as Appreciation.

*Really* (T23, T31): it was annotated as Engagement-Contraction and Graduation-Force. In the occurrence in text T23, it was annotated as Engagement-Contraction because it cannot be substituted by an intensifier like *very*, while it was precisely tagged as Contraction in T31 for that reason.

*Flojísima* [very-poor] (T43): it was annotated as Attitude-Judgement and Attitude-Appreciation. It was marked as Judgement because it refers to a novel's capability to accomplish its function (either entertaining, creating an interesting plot, surprising, etc.).

*Lenta* [slow] (T43, T44): it was annotated as Attitude-Judgement and Attitude-Appreciation. Like the previous case, it points to the novel's (in)capability to accomplish its goal. It is a slow novel, that is, it fails to achieve the expected pace of a standard plot.

*Kid-friendly* (T46): it was annotated as Attitude-Appreciation and Attitude-Judgement. This modifier addresses the ethical qualities of a film (respect, appropriate content, etc.), so it was tagged as Attitude-Judgement.

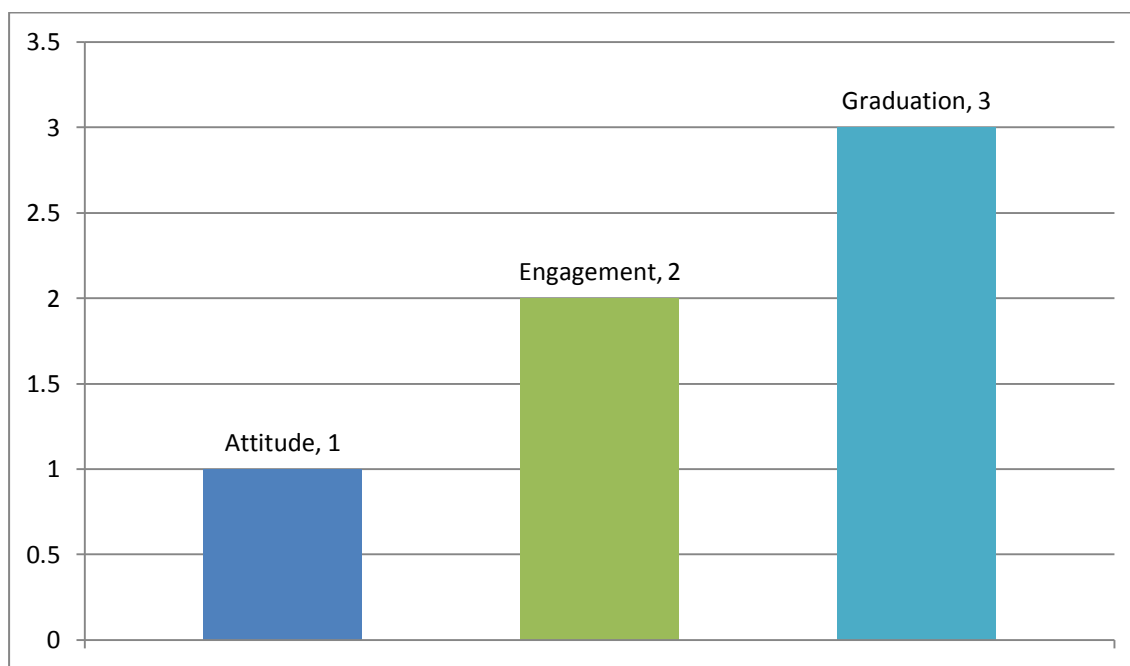
*Sweet* (T46): it was annotated as Attitude-Appreciation and Attitude-Judgement. The tag selected is the same as in the previous case, since it also evaluates a film's (metaphorical) attitude towards the public.

*Stunned* (T47): it was annotated as Attitude-Affect and Attitude-Appreciation. It refers to the feeling that the author experiences instead of an action attributed to an object, so the tag chosen was the former one.

*Definitely* (T48): it was annotated as Engagement-Contraction and Engagement-Expansion. The latter was chosen because it refers to a high level of certainty instead of emphasising the author's opinion

*Capital STUPID* (T50): it was annotated as Graduation-Focus and Graduation-Force. It is closer to the meaning of *very* than to the prototype of stupid, so the final tag was Graduation-Force.

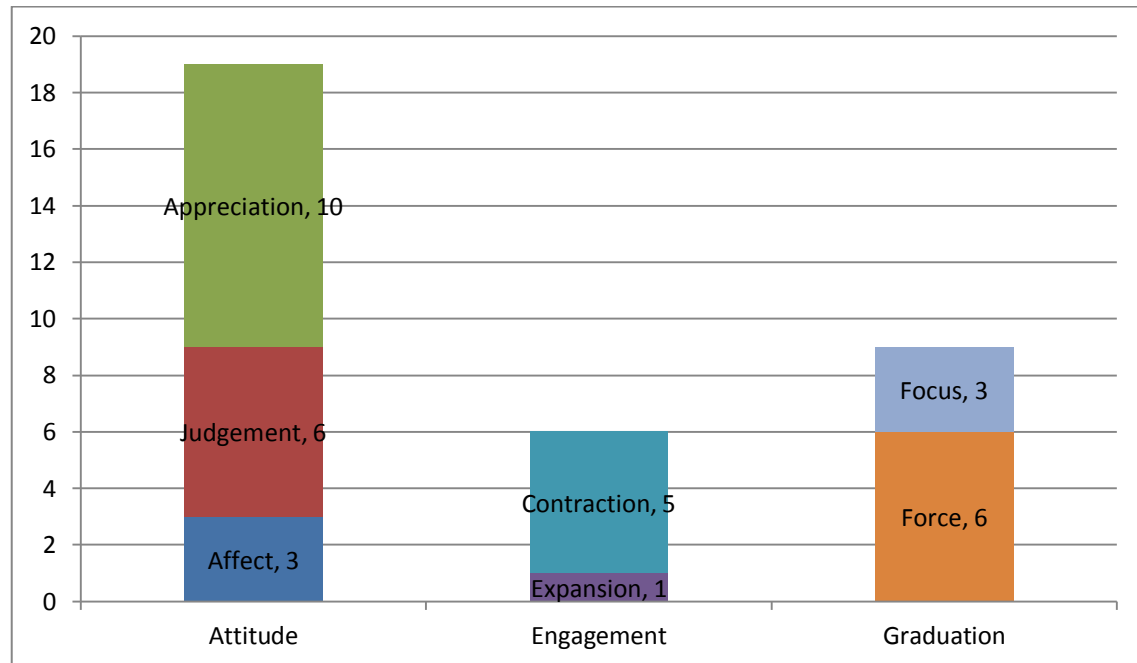
Figure 13 shows the distribution of disagreement when different major categories (Attitude, Engagement and Graduation) were selected by the annotators.



**Figure 13. Disagreement distribution of major categories**

Although it was not common, as the few occurrences show (1, 2 and 3 for Attitude, Engagement and Graduation, respectively), sometimes annotators did not agree on the major category the span belonged to. This means that a markable was annotated as Graduation when it was actually Attitude or Engagement, and the like. It was actually Graduation the category which was most frequently mismatched, followed by Engagement and Attitude.

Figure 14 shows the distribution of disagreement regarding major subcategories (Affect, Judgement, Appreciation, Expansion, Contraction, Force and Focus).



**Figure 14. Disagreement distribution of major subcategories**

In this case, the most inwardly-blurred category is Attitude: Appreciation has obtained the highest results in mismatching, while Judgement and Affect had lower results. However, it was mainly confused with other Attitude categories, which explains why Attitude had the lowest results in major categories disagreement. The second highest results can be seen in Graduation, with higher results for Force than Focus, followed closely by Engagement, where Contraction was clearly more confusing than Expansion.

Finally, Figure 15 shows the distribution of the most controversial combinations of annotations, that is, which two categories were typically used one instead of the other.

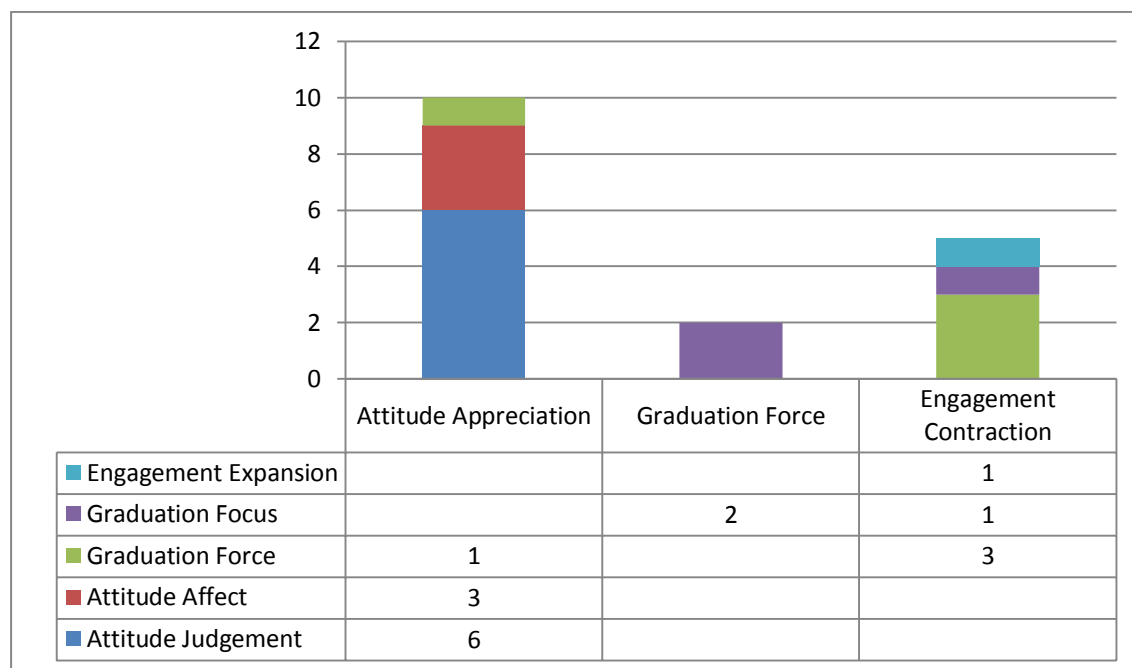


Figure 15. Disagreement distribution of categories combination

The most confused combinations are Attitude-Appreciation and Attitude-Judgement. They both refer to other elements than the author's feelings, and this may have caused the confusion. Theoretically, Judgement refers to other people's behaviours while Appreciation focuses on objects and natural phenomena. However, evaluative elements on moral aspects, typically used for human beings, can be associated with objects in a metaphorical way.

Attitude-Appreciation was also annotated as Attitude-Affect and vice versa. This is due to the fact that it is not clear when the focus is on the object causing a feeling or the author having that feeling caused by the object. Similar results were obtained by Graduation-Force and Engagement-Contraction. This is basically due to the word *really*, which has different meanings that are not always clearly distinguished.

Graduation-Force and Graduation-Focus had lower results. Their main problem with these combinations was related to elements like *that*, which did not clearly intensify a quality or a category.

Finally, other combinations that produced disagreement are Engagement-Contraction and Engagement-Expansion/Graduation-Focus on the one hand, and Attitude-Appreciation and Graduation-Force on the other hand. They had some marginal results whose individual mismatches were detailed above.



### 3.6.3. Annotation results for agreement study 3

The last agreement study asked the two annotators to select one of the most specific categories inside each of the Appraisal axes. Again, the previous step was agreed so that annotators could work on a common template with the same major categories already selected. Three k-values have been obtained, one for each of the axes, Attitude, Engagement and Graduation. Table 9 presents the results for the Attitude axis, including the categories Happiness, Unhappiness, Security, Insecurity, Satisfaction, Dissatisfaction, Normality, Capability, Tenacity, Veracity, Propriety, Reaction, Composition and Valuation.

		Annotator 1															
		Ha	Un	Se	In	Sa	Di	No	Ca	Te	Ve	Pr	Re	Co	Va		
Annotator 2	Ha	31	0	0	0	0	0	0	0	0	0	0	0	0	0	31	10%
	Un	0	9	0	0	0	0	0	0	0	0	0	0	0	0	9	3%
	Se	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
	In	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	0%
	Sa	0	0	0	0	31	0	0	0	0	0	0	0	0	0	31	10%
	Di	0	0	0	0	0	16	0	0	0	0	0	0	0	0	16	5%
	No	0	0	0	0	0	0	16	0	0	0	0	0	0	0	16	5%
	Ca	0	0	0	0	0	0	0	20	0	0	0	0	0	0	20	6%
	Te	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0%
	Ve	0	0	0	0	0	0	0	0	0	4	0	0	0	0	4	1%
	Pr	0	0	0	0	0	0	0	0	0	0	8	0	0	0	8	3%
	Re	0	0	0	0	0	0	0	0	0	0	0	48	0	0	48	15%
	Co	0	0	0	0	0	0	0	0	0	0	0	0	7	1	8	3%
	Va	0	0	0	0	0	0	0	0	0	0	0	6	1	115	122	39%
Tot		31	9	0	1	31	16	16	20	0	4	8	54	8	116	314	100%
		10%	3%	0%	0%	10%	5%	5%	6%	0%	1%	3%	17%	3%	37%	100%	

Table 9. Inter-annotator agreement in minor Attitude categories selection

The  $Pr(a)$  value is 0.60 and the  $Pr(e)$  is 0.20, obtaining a k-value of 0.49, a moderate agreement. This is the lowest result obtained in any k-value in this study. This means that the specific categories inside Attitude have present the highest complexity in the annotation process, and the boundaries between them were blurred. As the table shows, the most confusing categories were Reaction and Valuation, which were annotated differently on several occasions. The list of spans that caused disagreement is presented below.

*Inestable* [unstable] (T3): it was annotated as Normality and Capability. However, it does not refer to how usual it is in time, but to its capability to work as expected. The disagreement here was due to a different interpretation of the context.

*Sucks* (T9, T20): it was annotated as Valuation and Reaction. The successful tag was Valuation because it does not really imply causing a feeling to the speaker or other people, but it assigns a quality or value, in this case, negative, to the object addressed. Similar problems were found in *rocks* (T38) and *stunk* (T50), which were annotated both as Valuation and Reaction, and were finally classified as Valuation.

*Thinking* (T18): it was annotated as Valuation and Reaction. It refers to a type of game, but more specifically the game makes the player think. This was the reason why the tag chosen was Reaction.

*Pobre* [poor] (T22): it was annotated as Valuation and Reaction. Although it may seem that a value is assigned to an object or person, it is actually the reaction or feeling of piety that it triggers what made it be classified as Reaction.

*Absurd* (T28): it was annotated as Composition and Valuation. The key here lies in the issue it addresses. It refers to how badly a novel or part of it is structured. Therefore, Composition was the most adequate tag.

*Lovable* (T46): it was annotated as Valuation and Reaction. Instead of a quality of the thing itself, it is more predominantly perceived as the feeling it causes to the public, so it focuses on the reaction, the item makes you love it.

Table 10 presents the results for Engagement categories: Epistemic, Evidential, Pseudo-Question, Deontic, Acknowledge, Distance, Deny, Counter, Concur, Pronounce and Endorse.

		Annotator 1												
		Ep	Ev	PQ	D	Ack	Dis	De	Coun	Con	Pro	End		
Annotator 2	Ep	9	0	0	0	0	0	0	0	0	0	0	9	4%
	Ev	0	5	0	0	0	0	0	0	0	0	0	5	2%
	PQ	0	0	7	0	0	0	0	0	0	0	0	7	3%
	D	0	0	0	7	0	0	0	0	0	0	0	7	3%
	Ack	0	0	0	0	4	0	0	0	0	0	0	4	2%
	Dis	0	0	0	0	0	5	0	0	0	0	0	5	2%
	De	0	0	0	0	0	0	74	0	0	0	0	74	35%
	Coun	0	0	0	0	0	0	0	86	0	0	0	86	41%
	Con	0	0	0	0	0	0	0	0	1	1	0	2	1%
	Pro	0	0	0	0	0	0	0	0	0	11	0	11	5%
	End	0	0	0	0	0	0	0	0	0	0	0	0	0%
Total		9	5	7	7	4	5	74	86	1	12	0	210	100%
		4%	2%	3%	3%	2%	2%	35%	41%	0%	6%	0%		

Table 10. Inter-annotator agreement in minor Engagement categories selection

The  $Pr(a)$  value is 1 and the  $Pr(e)$  is 0.30, obtaining a k-value of 0.99, an almost perfect agreement. Actually, there is only one single case in which the annotators used different tags for the same span. The span that caused disagreement is the following one:

*Hay que reconocer* [it must be admitted] (T22): it was annotated as Pronounce and Concur. Both meanings are very similar since they do only differ in subtle nuances regarding the emphasis on the author's authorial voice as basis for the statement or clear general evidence that makes the statement undeniable. The tag selected was Concur, since the authorial voice was not strong enough.

Finally, Table 11 presents the results for Graduation categories: Isolation, Infusion, Repetition, Mass, Number, Extension, Sharpen and Soften.

		Annotator 1									
		Iso	Inf	Rep	Mass	Num	Ext	Sharp	Soft		
Annotator 2	Iso	89	0	0	0	0	0	0	0	89	44%
	Inf	0	63	0	0	0	0	0	0	63	31%
	Rep	0	1	7	0	0	0	0	0	8	4%
	Mass	0	0	0	1	0	0	0	0	1	0%
	Num	0	0	0	0	36	0	0	0	36	18%
	Ext	0	0	0	0	0	0	0	0	0	0%
	Sharp	0	0	0	0	0	0	3	0	3	1%
	Soft	0	0	0	0	0	0	0	1	1	0%
		89	64	7	1	36	0	3	1	201	100%
		44%	32%	3%	0%	18%	0%	0%	0%	100%	

Table 11. Inter-annotator agreement in minor Graduation categories selection

The  $Pr(a)$  value is 1 and the  $Pr(e)$  is 0.33, obtaining a k-value of 0.99, an almost perfect agreement. Just like in the case of Engagement, there is only one span that was tagged in different ways by the annotators. The span that caused disagreement is the following one:

*Peor* [worst] (T5): it was annotated as Repetition and Infusion. The choice of Repetition was due to the occurrence of the second *peor* [worst] in the expression *lo peor de lo peor* [the worst in the worst]. However, it was definitely tagged as Infusion because it stands for the delimiting of a group (the worst one) and the repetition does not emphasise that meaning strongly enough.

Figure 16 shows the distribution of disagreement in the minor categories analysed in this section.

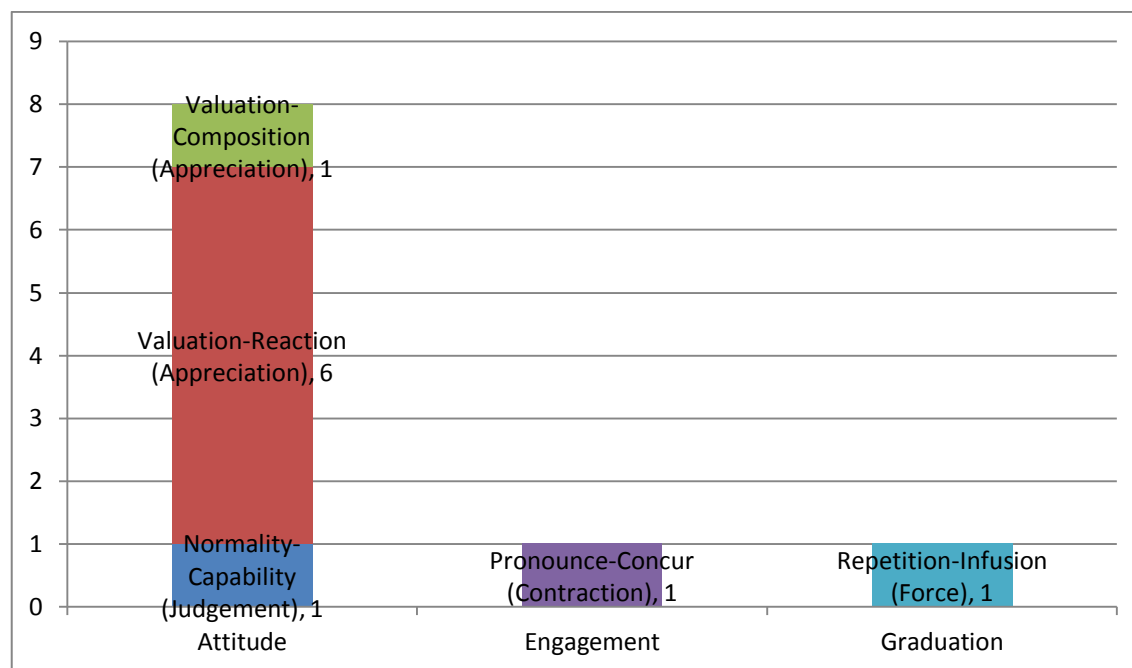


Figure 16. Disagreement distribution of minor categories

As can be clearly seen, the highest disagreement was found between the categories Reaction and Valuation. Reaction stands for the impact the object has on the speaker, while Valuation refers to how worthwhile or valuable the object is. These two categories can be easily mixed due to interpretation issues since one can be seen as causing the other. The line that places the emphasis on one or the other is sometimes in the eye of the beholder.

### 3.7. Summary and discussion

This chapter has described the methodological steps which have been taken to use corpus annotation for the validation of a number of aspects of Appraisal Theory empirically, following the methodology proposed by Hovy and Lavid (2010), and Lavid (2012). First, the training corpus of 50 texts was presented; secondly, the annotation scheme and procedure were described; and, finally, the results of the agreement studies were described in detail showing the areas and aspects which caused disagreement between coders.

The Kappa coefficient showed that a high value of agreement was reached in the annotation of the training corpus as a whole. This means that the annotations were highly

reliable and the corpus could be consequently used as a resource for further linguistic analyses and as a training corpus for computational purposes. The lowest results were obtained when tagging minor Attitude categories, more specifically, Reaction and Valuation. Both Engagement and Graduation categories obtained an exceptionally high  $k$ -value, very close to 1, so categories in these two axes can be clearly distinguished.

The detailed listing of the realisations exemplified some confusing cases where different interpretations can be obtained by different coders, even when following the same instructions. These mismatches can be due to either the ambiguity of the theoretical categories themselves or to an insufficient description or exemplification in the annotation guidelines.

Regarding span selection, most disagreements were related to long phrases and sentences implying an evaluative meaning beyond the mere description of actions. With respect to coarse categories, Graduation appeared as the most conflictive category, which points to an unclear difference between intensification or additional description and values. Subcategories in Attitude (Affect, Judgement and Appreciation) showed the highest intersections, showing that even though Attitude is a category that can be clearly identified, subtypes within this category are not so noticeably different. This is closely related to the use of personifications (assignation of personal qualities to objects). Consequently, minor categories inside Appreciation, more specifically, Valuation and Reaction, showed the highest variation between coders, where the use of terms other than adjectives made it more difficult to distinguish objects' qualities (Valuation) from objects as sources of others' qualities or emotions (Reaction).

## Chapter 4

# ANNOTATION OF THE LARGER CORPUS

### 4.1. Introduction

Following the methodology proposed in Hovy and Lavid (2010), the next step in the annotation methodology is to annotate a larger dataset with the validated tags of the proposed annotation scheme. This larger dataset —also called the ‘test’ corpus in NLP— was manually annotated by the author of this dissertation using the tags of the annotation scheme validated in chapter 3. For this task, I followed Dligach et al. (2010) suggestion, according to which “instead of double annotating and adjudicating the common accepted wisdom in annotation projects, it is often better to single annotate more data because it is a more cost-effective way to achieve a higher performance” (Dligach et al. 2010).

The purpose of this annotation task was to investigate the research questions (2), (3) and (4) presented in the introduction, namely:

- Are there language-independent Appraisal features which characterise mobile application reviews in English and Spanish? If so, which are these features?
- Are there language-specific Appraisal features of mobile application reviews written in English and Spanish?

- Are there significant differences in the use of Appraisal features depending on the type of item reviewed and depending on the polarity features of the review?

In order to address these questions, the larger corpus will be presented and English and Spanish datasets will be described and compared in detail. The results are presented according to three factors: language-specific differences, item differences (application, game, book, film or music), and polarity differences, i.e., comments rated as positive (5 or 4 stars) or negative (1 or 2 stars).

## 4.2. The larger corpus

The larger corpus used in this study consists of a collection of 200 texts filtered and selected following the same procedure as the training set. This means they include comparable English and Spanish texts distributed evenly as illustrated in Tables 12 and 13 (st. stands for ‘stars’, regarding the 1-to-5 star rating).

Applications 20				Games 20				Books 20				Films 20				Music 20			
+	-			+	-			+	-			+	-			+	-		
10	10			10	10			10	10			10	10			10	10		
5st.	5	1st.	5	5st.	5	1st.	5	5st.	5	1st.	5	5st.	5	1st.	5	5st.	5	1st.	5
4st.	5	2st.	5	4st.	5	2st.	5	4st.	5	2st.	5	4st.	5	2st.	5	4st.	5	2st.	5

Table 12. Selection of 100 English texts for the larger dataset

Applications 20				Games 20				Books 20				Films 20				Music 20			
+	-			+	-			+	-			+	-			+	-		
10	10			10	10			10	10			10	10			10	10		
5st.	5	1st.	5	5st.	5	1st.	5	5st.	5	1st.	5	5st.	5	1st.	5	5st.	5	1st.	5
4st.	5	2st.	5	4st.	5	2st.	5	4st.	5	2st.	5	4st.	5	2st.	5	4st.	5	2st.	5

Table 13. Selection of 100 Spanish texts for the larger dataset

The texts addressed several items inside each of the products. This decision was made in order to enhance diversity in the texts inasmuch as authors commenting on two different applications may cover more issues than those commenting on the same one.



The comments in the test corpus address at least two items per category. These items are the following ones:

6. Clean Master (application): an application that frees up some space in your smartphone.
7. Instagram (application): a social networking application that allows users to upload images and videos.
8. Angry Birds (game): a social networking application that allows users to upload images and videos.
9. Candy Crush (game): a match-three puzzle game.
10. *All the Light We Cannot See* (book): a historical novel by Anthony Doerr, published in 2014.
11. *Fifty Shades of Grey* (book): an erotic romantic novel by E. L. James.
12. *The Girl on the Train* (book): a psychological thriller novel by Paula Hawkins, published in 2015.
13. *Avatar* (film): an epic science-fiction film directed by James Cameron in 2009.
14. *Gravity* (film): an adventure film set in space directed by Alfonso Cuarón in 2013.
15. *Frozen* (film): a 3D computer-animated musical fantasy film produced by Disney.
16. *The Wolf of Wall Street* (film): a biographical crime comedy film directed by Martin Scorsese in 2013.
17. AC/DC (music): a rock and heavy metal music group.
18. Adele (music): a soul, pop and R&B singer.

The annotation environment used in this part of the study was the UAM Corpus tool<sup>5</sup>. The UAM Corpus Tool is a free state-of-the-art environment for annotation of text corpora. It supports annotation of multiple texts at multiple linguistic levels (Clause, Sentence, Document, etc.) as well as analysis methods such as instances retrieval and statistical measurements. Figure 17 shows UAM Corpus Tool home screen.

---

<sup>5</sup> UAM Corpus Tool can be downloaded for free at [www.corpustool.com](http://www.corpustool.com). The version used in this study was 3.3.

#### 4. Annotation of the larger corpus

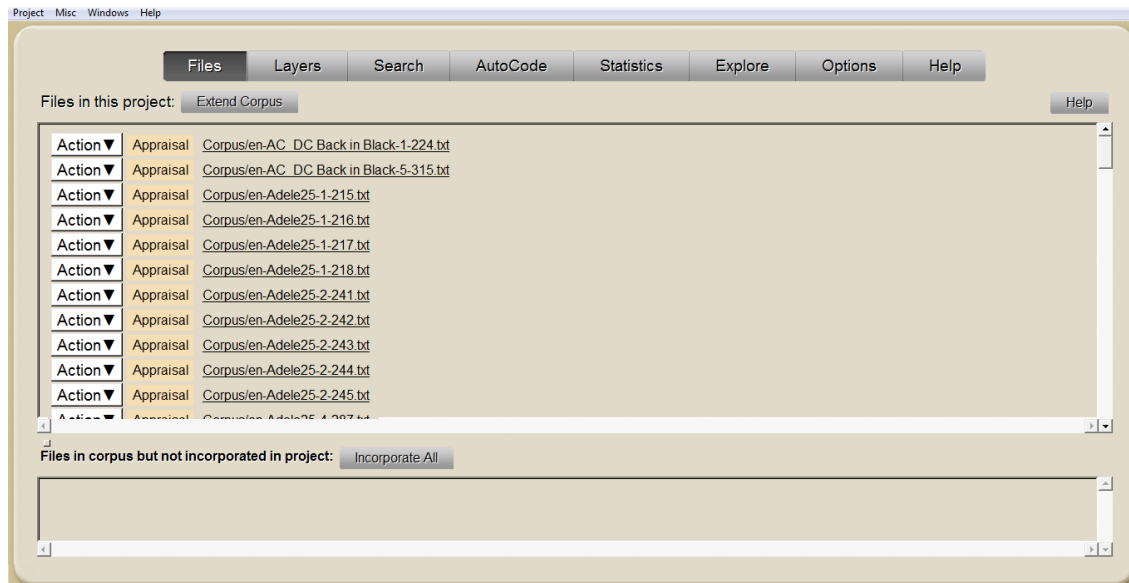


Figure 17. UAM Corpus Tool home screen

It was the most adequate tool for this study since it does not include more options than needed, which makes it simpler to use, and is very user-friendly. Figure 18 illustrates the annotation interface in UAM Corpus Tool.

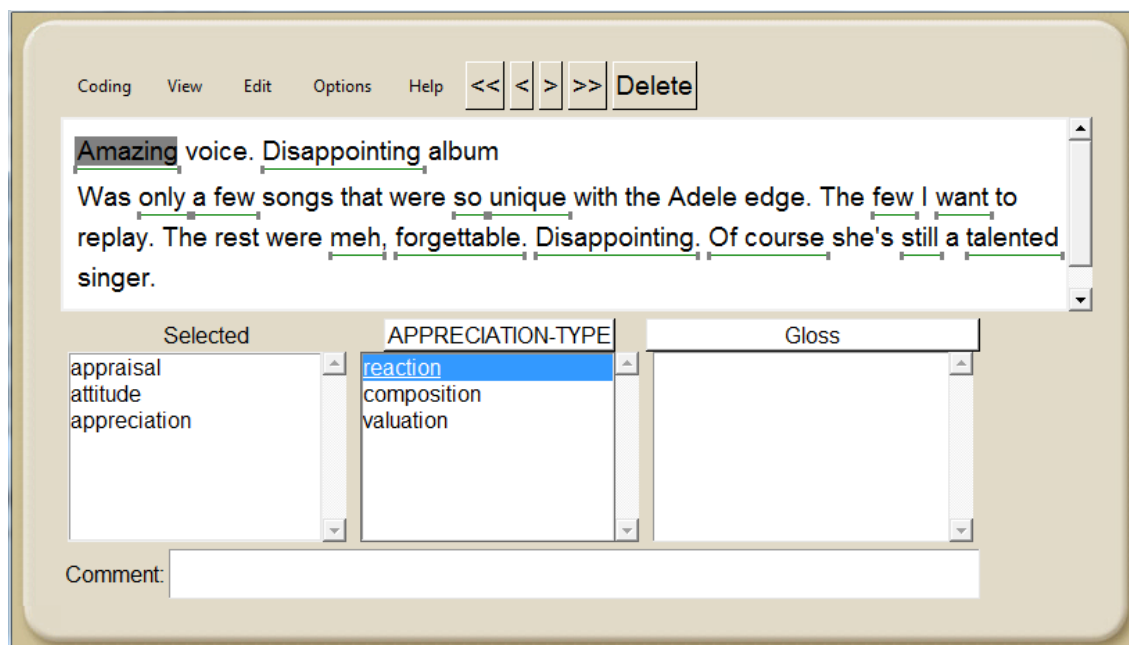


Figure 18. UAM Corpus Tool annotation screen

Unlike in the analysis of the training corpus, texts did not require a prior division into sentences or other smaller units. Texts are uploaded to the tool for their annotation and

spans in the text are manually selected and assigned a tag previously included in a layer or schema. Thus, the same corpus can be annotated following different schemata which can be compared at a later stage. In this case, only the Appraisal layer was used, including the already mentioned categories.

Sections 4.3, 4.4, 4.5. and 4.6. below present the annotation results. For the sake of convenience, a summary of the core and the extended Appraisal tagsets used in the study are included here again.

Attitude	Feelings, including emotional reactions, judgements of behaviour and evaluation of qualities of things.
Affect	Emotional reactions and feelings
Judgement	Assessment of behaviour according to normative principles
Appreciation	Evaluation and valuation of things
Engagement	Implication of other possibilities and voices than the speakers'
Expansion	Author's position is one inside a range of possible options or an external source provided for a given opinion
Contraction	Author positions against a contrary position or limits the scope of possibilities
Graduation	Grading phenomena whereby feelings are amplified or softened and categories blurred or sharpened.
Focus	Degree of prototypicality
Force	Degree of intensity or amount

**Table 14. Core tagset for Appraisal annotation**

Happiness	Feelings of cheer and affection
Unhappiness	Feelings of misery and antipathy
Security	Feelings of confidence and trust
Insecurity	Feelings of disquiet and surprise
Satisfaction	Feelings of interest and pleasure
Dissatisfaction	Feelings of ennui and displeasure
Normality	How (un)usual someone or something is
Capacity	How (in)capable someone or something is
Tenacity	How (ir)resolute someone or something is
Veracity	How truthful someone or something is
Propriety	How ethical someone or something is
Reaction	Effect, impact and appearance
Composition	Balance and complexity
Valuation	Value and worthiness
Epistemic	Degree of certainty
Evidential	Empirical evidence
Pseudo-Question	Questions with unintended answer
Deontic	Obligation, control and compliance
Acknowledge	External source provided for a given opinion with no overt indication as to where the authorial voice stands
Distance	External source provided for a given opinion with an explicit distancing of the authorial
Deny	Negation
Counter	Concession and counter expectation
Concur	Overt announcement of agreement or shared knowledge with a projected alternative
Pronounce	Authorial emphases
Endorse	External sources construed by the authorial voice as correct and undeniable
Sharpen	Maximisation of the degree of prototypicality
Soften	Lessening of the degree of prototypicality
Isolation	Scaling of intensity realised by an individual lexical item
Infusion	Scaling of intensity realised by one aspect of the meaning of a single term
Repetition	Repetition of the same lexical item or part of it
Number	Numerical quantifiers
Mass	Extension of an item's presence in space
Extent	Proximity and distribution in time and space

**Table 15. Extended tagset for Appraisal annotation**

### 4.3. Annotation results: general tendencies

Before starting with the contrastive analysis, an account of the annotation results of the corpus as a whole is displayed in Figure 19.

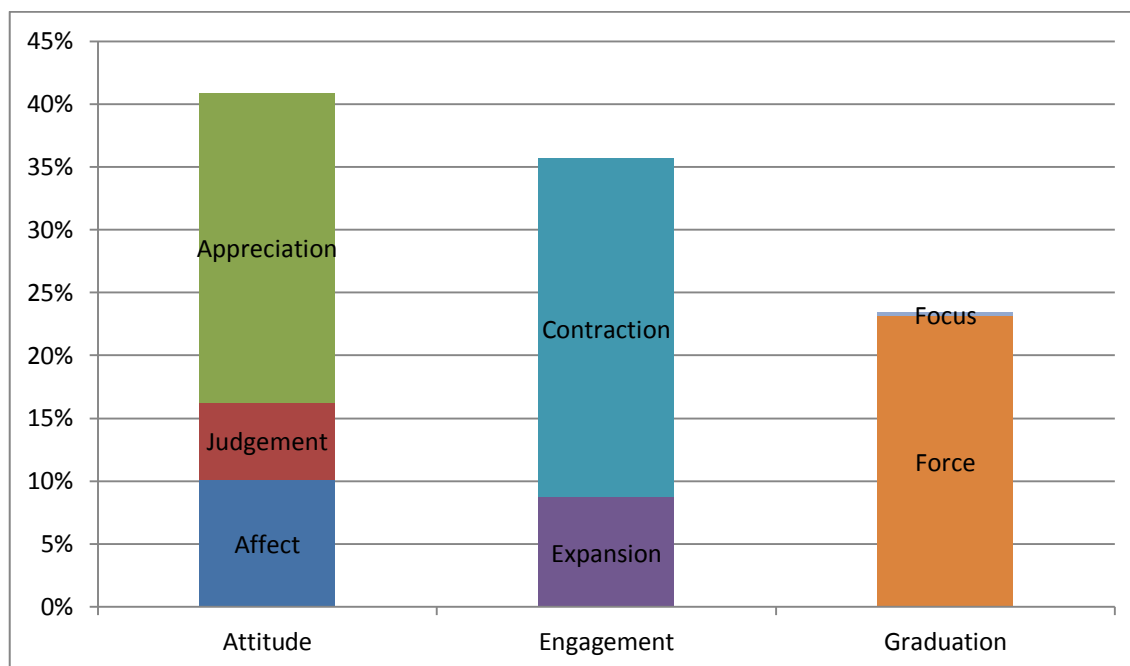


Figure 19. Distribution of occurrences per category in the larger dataset

The graph illustrates the distribution of occurrences in percentages considering the number of spans which were selected and coded with a specific tag in relation to the total number of spans marked in the whole corpus (3086 spans, 1766 in English and 1320 in Spanish).

At a general level, the most frequently annotated category was Attitude (40.89%), followed by Engagement (35.64%) and Graduation (23.46%). However, when looking at the more specific tags, the most frequent one was Contraction (26.93%). This is due to the number of negations (Deny) and hypothetical situations (Counter) that are included in both languages. The second most common category was Appreciation (24.59%), which should be expected since the annotated texts are rich in Valuation or expressions conveying a value associated with an object and their aim is to describe those reviewed items. Finally, the tag with the third highest number of occurrences is Force (23.14%),

which includes all those intensifiers and quantifiers that increase or lower the value of, mainly, other nouns, adjectives and verbs.

Regarding the lowest results, Focus (0.32%) obtained the scarcest representation. The annotated markables in this category are used to soften or sharpen the boundaries of a word, that is, to express how close it is to the prototypical idea of that item, but users prefer to quantify nouns than stress or diminish their core meanings. Judgement (6.19%) is used to assign social or moral values to people. However, it will be exemplified how this is not only used to address people but also objects. In any case, this kind of value was not a pivotal one in the items selected. Thirdly, Expansion (8.72%) occurrences or expressions showing different levels of certainty and allowing for other opinions apart from the authorial one obtained less than half the results of their counterpart Contraction. This means that reviewers place the stress on their own voice, limiting the possibilities of other options, instead of presenting their opinion as one of a range of possible choices. Finally, Affect (10.11%) is placed in the very middle of the ranking. This type of expressions refers to someone's feelings. Therefore, Affect expressions are used to describe how the author (or other users) feel with respect to the item reviewed and, in spite of their occurrence in the annotated texts, comments focus much more on the value or even the effects of the item itself than on users' feelings.

Some words or expressions were frequently shared by different texts. In Table 16, the ten most common spans are presented, arranged by the major Appraisal category they belong to as well as the number of occurrences of each and the percentage they represent in relation to the total number of Appraisal annotated items.

Attitude			Engagement			Graduation		
#	%	Span	#	%	Span	#	%	Span
27	2.14	love	135	12.27	no	47	6.49	!
24	1.90	great	85	7.73	but	42	5.80	muy [very]
21	1.66	good	59	5.36	pero [but]	26	3.59	!!
17	1.35	bien [well]	54	4.91	not	25	3.45	so
16	1.27	mejor	34	3.09	just	20	2.76	more
16	1.27	buena [good]	33	3.00	if	19	2.62	most
14	1.11	like	25	2.27	si [if]	18	2.49	very
13	1.03	awesome	19	1.73	only	16	2.21	!!!
12	0.95	fun	16	1.45	don't	16	2.21	más [more]
12	0.95	amazing	16	1.36	even	16	2.21	mejor [better, best]

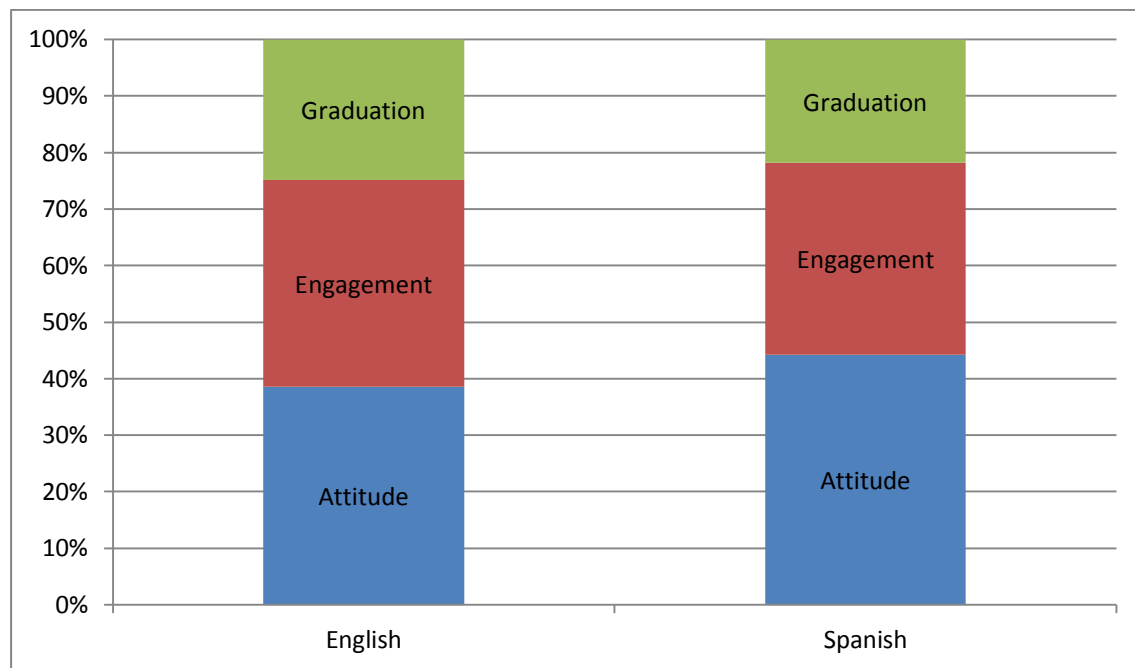
Table 16. Most common Appraisal spans in Attitude, Engagement and Graduation

Attitude spans are basically verbs and adjectives: verbs include *love* and *like*, while adjectives are *great*, *good*, *mejor*, *buena* [good], *awesome* and *amazing*, and only one adverb was found, the Spanish *bien* [well]. Engagement spans are typically conjunctions and adverbs: conjunctions include *but*, *pero* [but], *if* and *si* [if], while the adverbs are *no* (the same word for both English and Spanish, which explains the high number of marked spans), *not*, *just*, *only* and *even*, and an additional short form used for negation, *don't*. Finally, spans in Graduation are either exclamation marks (!), which can be typed in a row (!! , !!!), adverbs, such as *muy* [very], *so*, *very*, and *más* [more], or comparative and superlative forms, such as *most*, and *mejor* [better, best].

This initial overview of the data has provided a glimpse of the general tendencies in the annotation of Appraisal features in the whole corpus. In the following sections, I will describe the annotation results, according to the criteria of language, item, and polarity.

#### 4.4. Language-specific differences (English vs. Spanish)

The first of the three criteria used is language, that is, the 100 English reviews are compared with the 100 Spanish reviews, disregarding the item reviewed or the polarity (positive or negative) of the comment. Figure 20 shows the distribution of Attitude, Engagement and Graduation in English and Spanish.



**Figure 20. Distribution of Appraisal features in English and Spanish datasets**

As a first approach, English shows a slightly higher preference for Engagement (36.92%) than Spanish (33.94%), as well as for Graduation (24.69% in English vs. 21.82% in Spanish). However, the most visible difference involves Attitude, where it was found that Spanish occurrences go up to 44.24% while only a 38.39% of the English tags are marked as Attitude.

This means that Spanish authors are more prone to include expressions about feelings and evaluations while English authors use a higher number of modulators, either including or denying others' opinions or intensifying or softening their language.

Regarding the use of specific expressions, Table 17 displays the ten most used expressions in each language.



English			Spanish		
#	%	Span	#	%	Span
85	4.81	but	127	9.62	no [no, not]
56	3.17	not	59	4.47	pero [but]
35	1.98	just	42	3.18	muy [very]
35	1.98	!	32	2.42	mejor [better, best]
34	1.93	if	25	1.89	si [if]
28	1.59	love	21	1.59	!!
27	1.53	so	20	1.52	bien [well]
26	1.47	really	17	1.29	buena [good]
24	1.36	great	16	1.21	más [more]
21	1.19	good	14	1.06	mas <sup>6</sup> [more]

Table 17. Most common Appraisal spans in English and Spanish

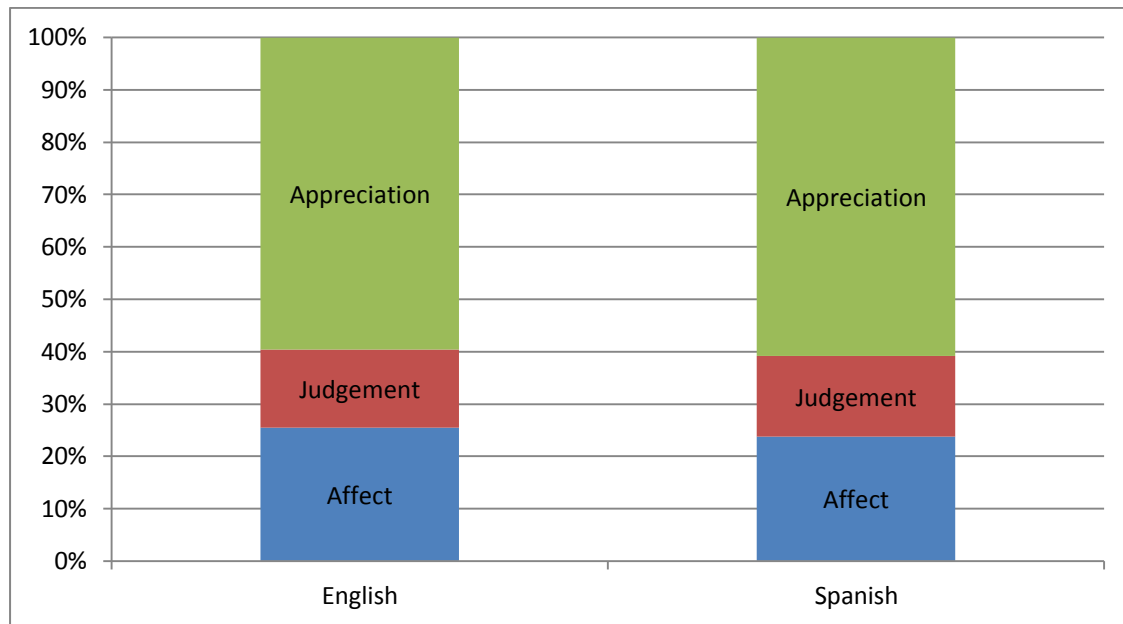
Very similar results are found in both languages regarding the type of words. Spans like *but*, *not*, *if* or *good* are among the most used ones in both languages. Differences include the use of double exclamation mark in the case of Spanish and single exclamation mark in English, showing that Spanish texts tend to be more emphatic. No verb is included in the Spanish ranking, while *love* does in English. English authors present a higher focus on the action while the Spanish ones draw on modifiers such as *más* [more] or *muy* [very].

These are only some initial contrastive aspects. The following subsections will analyse in detail the differences found in English and Spanish arranged by the Appraisal axis.

#### 4.4.1. Attitude

Attitude covers meanings about feelings, social judgement and objects valuations. Figure 21 below summarises the occurrences marked for its three major categories: Affect, Judgement and Appreciation.

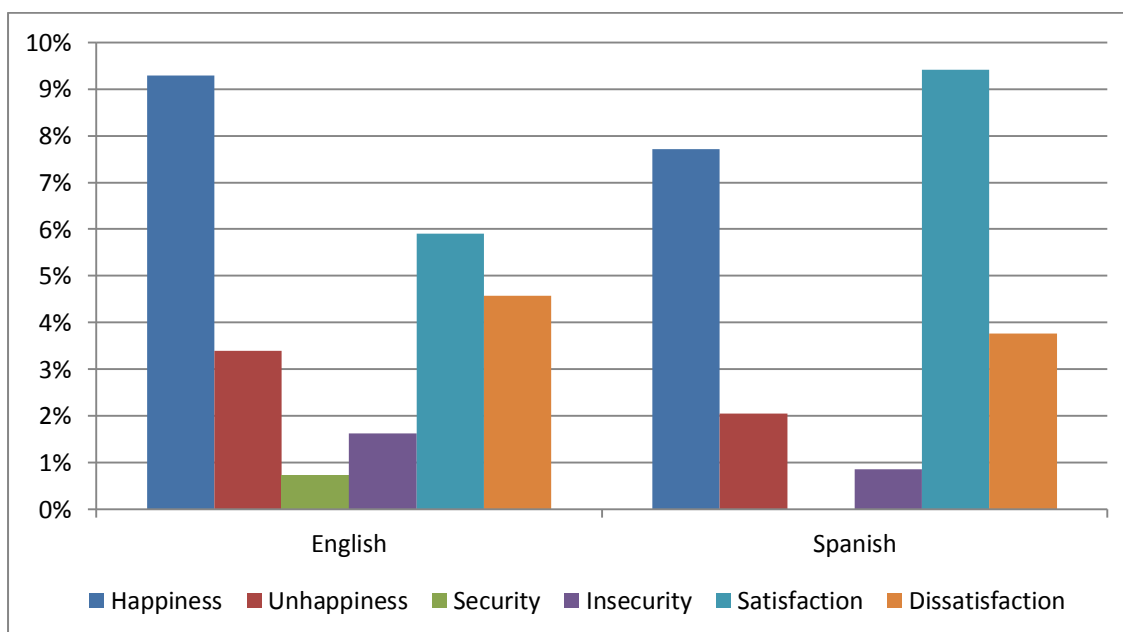
<sup>6</sup> The word *mas* in Spanish is a synonym of the conjunction *but*. However, the use found in the corpus is never this one, but a misspelling of *más* [more].



**Figure 21. Distribution of Attitude in English and Spanish**

A first look at the Attitude results shows no apparent difference between the two languages. Affect (25.52% vs. 23.80%), Judgement (14.90% vs. 15.41%) and Appreciation (59.59 vs. 60.79%) have almost equivalent results in English and Spanish respectively. However, a closer look at each of the subtypes of this axis will shed some light on divergent spots.

Affect is the first of the three categories analysed. Its results are illustrated in Figure 22.



**Figure 22. Distribution of Affect in English and Spanish**

The most visible differences in Affect involve Happiness and Satisfaction. While Happiness is more frequently used in English (9.29% vs. 7.71%) with verbs like *love* and *like*, Spanish reviewers turn to Satisfaction verbs (5.90% vs. 9.42%) that show interest, such as *querer* [want] and *esperar* [hope]. The rest of categories obtained close percentages: Unhappiness (3.39% vs. 2.05%), Security (0.74% vs. 0.00%), Insecurity (1.62% vs. 0.86%) and Dissatisfaction (4.57% vs. 3.77%).

It must be pointed out that the positive aspects Happiness and Satisfaction usually occurred more often than their negative counterparts, while Insecurity exceeded Security results. Actually, no span was tagged as Security in Spanish at all. This means that authors in both languages prefer a positive lexicon and the use of separate negation in case they want to deny those aspects, that is, a sentence like *I don't like it* is more likely to appear than *I dislike it*, and the like.

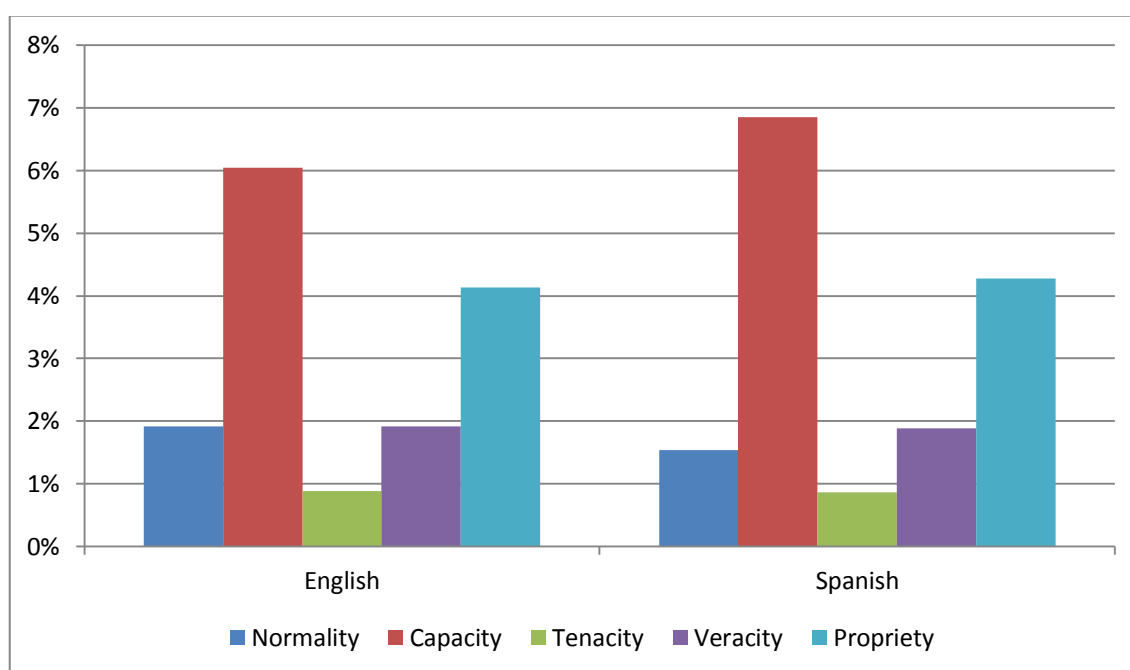


Figure 23. Distribution of Judgement in English and Spanish

Judgement is the subtype of Attitude that presents the fewer number of differences between the two languages. The percentages obtained are almost identical: Normality (1.92% vs. 1.54%), Capacity (6.05% vs. 6.85%), Tenacity (0.88% vs. 0.86%), Veracity (1.92% vs. 0.86%) and Propriety (4.13% vs. 4.28%).

Inside this category, Capacity and Propriety are the most recurrent types, with realisations like *strong*, *dumb*, *talented*, *estúpido* [stupid], *idiotas* [idiots] or *ingenuos* [naïve] for Capacity, and *racist*, *demonic*, *alcoholic*, *plagio* [plagiarism], *machista* [sexist] or *engañoso* [tricky] for Propriety. It must be noted that while Capacity spans covered positive and negative terminology, most terms tagged as Propriety had negative connotations, as the examples have shown.

Finally, Appreciation results are displayed in Figure 24.

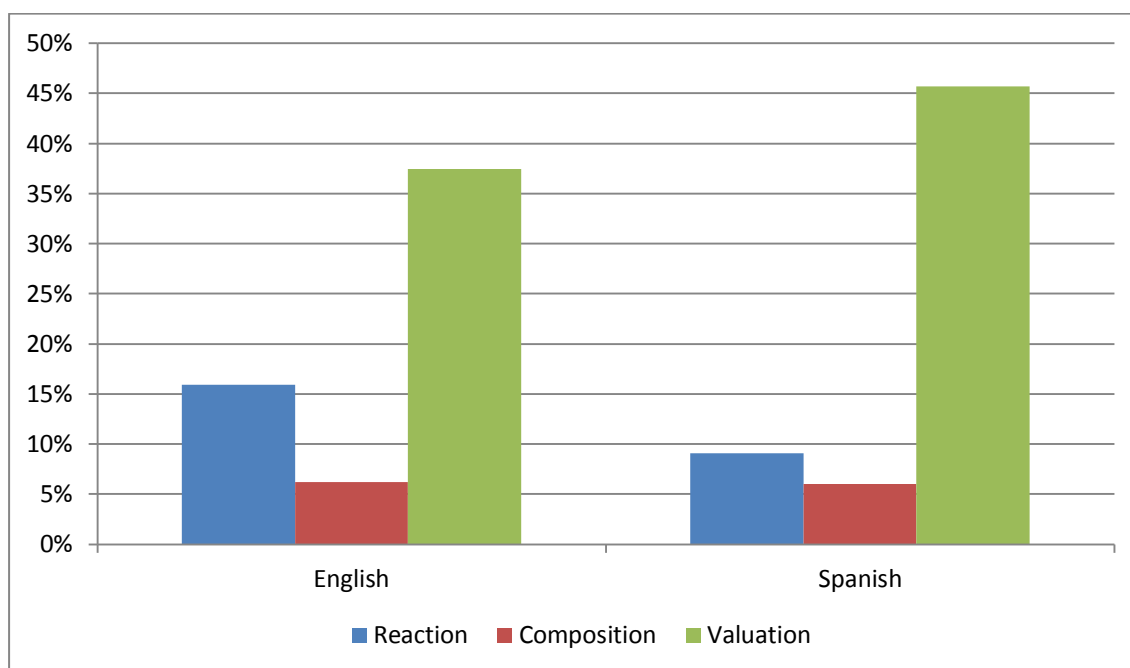


Figure 24. Distribution of Appreciation in English and Spanish

A much higher difference is observed in this category. The spans tagged as Reaction in English (15.93%) are almost twice the Spanish ones (9.08%), while the Appreciation spans have a higher representation in Spanish (45.72%) than in English (37.46%). English speakers focus on the effects of objects, which is shown in words like *amazing*, *fun*, *addictive* or *boring*, while Spanish speakers pay more attention to objects qualities such as *buena* [good], *mejor* [better, best], *excelent* [excellent] or *perfecta* [perfect].

Nevertheless, the results of Composition were virtually identical in both languages (6.19% vs. 5.99%) and typically referred to the level of complexity of the item analysed (e.g. *simple*, *intricate*, *balanced*, *difíciles* [hard], *sencillo* [simple], *fácil* [easy]).

Although some of the lexicon used has already been introduced, the most common spans used inside the Attitude axis are presented in Table 18.

English			Spanish		
#	%	Span	#	%	Span
27	3.98	love	17	2.91	bien [well]
24	3.54	great	16	2.74	mejor [better, best]
21	3.10	good	16	2.74	buena [good]
14	2.06	like	10	1.71	bueno [good]
13	1.92	awesome	9	1.54	recomiendo [(I) recommend]
12	1.77	amazing	8	1.37	5 estrellas [5 stars]
12	1.77	fun	8	1.37	lento [slow]
10	1.47	want	7	1.20	gusta [like]
9	1.33	best	7	1.20	quiero [want]
9	1.33	beautiful	6	1.03	genial [great]

Table 18. Most common Attitude spans in English and Spanish

Shared spans include *great&genial*, *good&buena* (feminine) / *bueno* (masculine), *like&gusta*, *want&quiero* and *best&mejor* (although *mejor* also stands for *better*). English authors also use frequently words like the verb *love* or adjectives like *awesome*, *amazing*, *fun* and *beautiful*, focusing mainly on the Reaction type, as mentioned above, while the Spanish ones prefer realisations such as *bien* [well], *recomiendo* [(I) recommend], *5 estrellas* [5 stars] and *lento* [slow], which cover tags like Valuation, Composition and Satisfaction, again distancing from feelings or assigning an effect to objects.

#### 4.4.2. Engagement

Engagement refers to expressions including the author's voice or other voices in the text. Figure 25 below presents the overview of the results for this second axis in both English and Spanish texts.

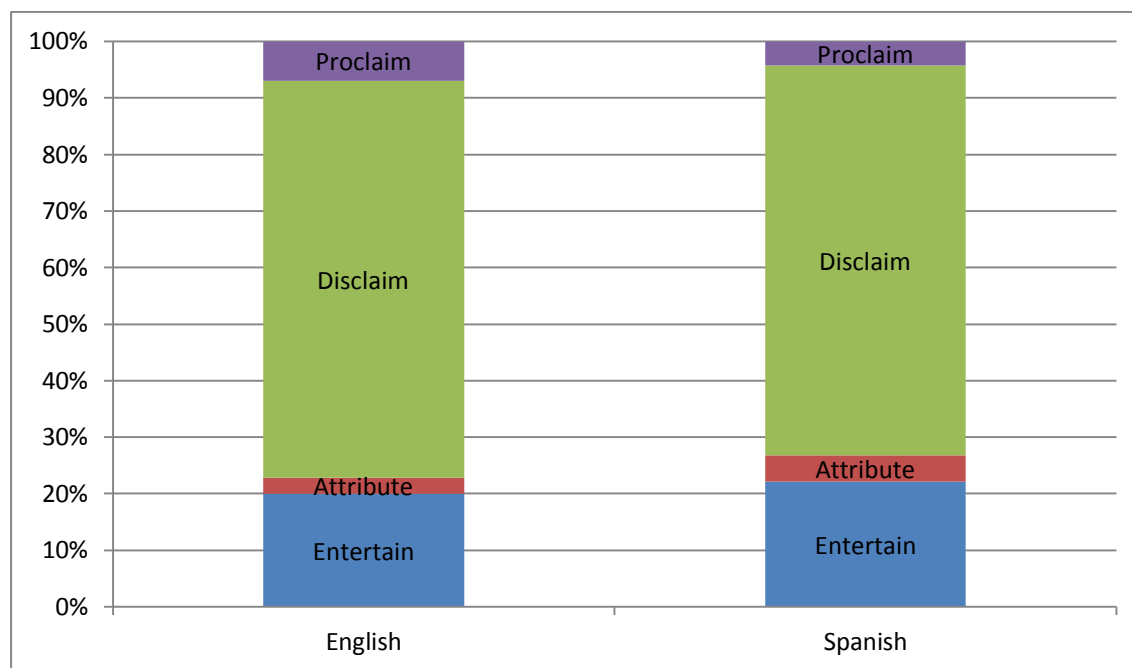


Figure 25. Distribution of Engagement in English and Spanish

The results may seem similar to some extent although there are some slight differences. These include higher results for Proclaim in English (6.90% vs. 4.24%) and lower results for Entertain (19.94% vs. 22.10%). Less distance was obtained in the percentages of Attribute (2.91% vs. 4.69%) and Disclaim (70.25% vs. 68.97). More important is the broad representation that the category Disclaim has as a member of Engagement in both languages, although a closer look at it will reveal further disparity.

Overall, the Expansion type, covering Entertain (Epistemic, Evidential, Pseudo-Question and Deontic) and Attribute (Acknowledge and Distance), and oriented to place the author's voice as just one among a range of other possible voices, is visibly less frequent than its counterpart, Contraction, including Disclaim (Deny and Counter) and Proclaim (Concur, Pronounce, Endorse), which aims to highlight one voice, either the author's or some else's as the correct one. Figure 26 gathers the results for the first category, Expansion, and its minor categories in both languages.

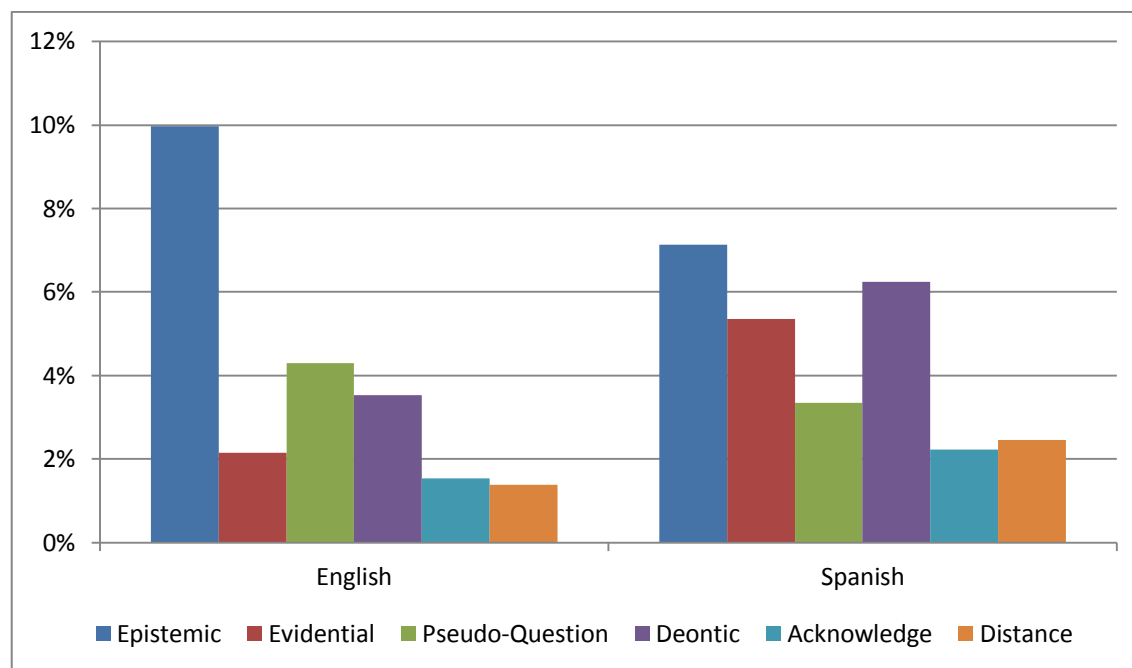


Figure 26. Distribution of Expansion in English and Spanish

The figure shows a clear divergence in most of Entertain categories: Epistemic, Evidential and Deontic. More specifically, English shows a higher preference for Epistemic spans (9.97% vs. 7.14%), using expressions that soften the authority and certainty of their voice and statements, such as *definitely*, *maybe*, *I thought*, *may*, etc.. On the other hand, Spanish Deontic spans are more recurrent (3.53% vs. 6.25%), indicating obligation of some kind (either something the author or some else must or has to do), in examples like *tienes que* [(you) have to], *debería* [(I/He/She/It) should] or *debéis* [(you-all) must]. Moreover, Spanish authors use Evidential markers more often (2.15% vs. 5.36%), linking their opinions to some visible evidence, as in *parece* [(he/she/it) seems], *se supone* [(it) is supposed], *parecía* [(I/he/she/it) seemed]. These three qualities together point to a more aggressive or strong speech on the Spanish side than on the English one.

In addition, English texts do also manifest a more frequent use of Pseudo-Questions (4.29% vs. 3.35%) or questions that do not look for an explicit answer (*how can this be romantic*, *what can I say*, *where did they learn to do that*, etc.), while the Spanish ones use slightly more recurrently Acknowledge (1.53% vs. 2.23%) and Distance (1.38% vs. 2.46%) expressions (*sabes* [(you) know], *dices* [(you) say], etc.), that is, they quote or refer to other voices more often than can be observed in the English texts. Therefore, English texts show feelings of surprise and indignation by means of

introducing questions in their comments, questions which do not actually want to be answered, while Spanish texts refer to common knowledge to support their own opinions.

The use of expressions constraining other textual voices is displayed in Figure 27. It includes Disclaim (Deny and Counter) and Proclaim (Concur, Pronounce and Endorse) categories.

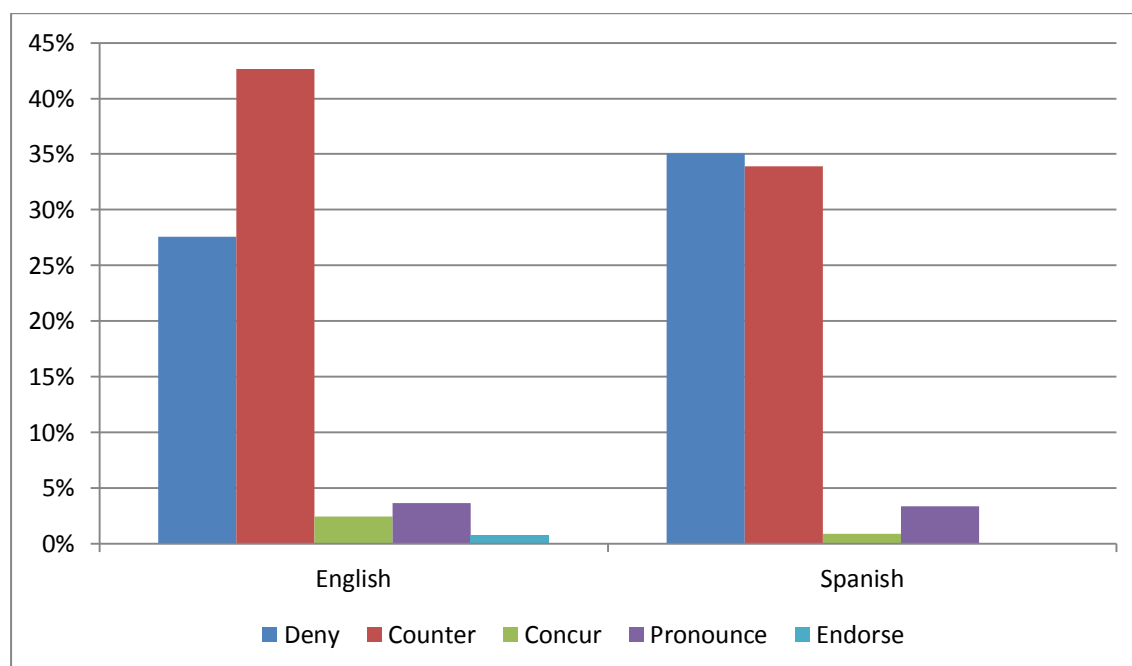


Figure 27. Distribution of Contraction in English and Spanish

The figure visibly shows an overwhelming presence of Deny and Counter elements in both languages in detrimental to Concur, Pronounce and Endorse. However, the results for Deny and Counter in English are very different from the Spanish ones. In the first place, the presence of Deny in English is lower than the same category in Spanish (27.61% vs. 35.04%), while the latter is less profuse in Counter (42.64% vs. 33.93%). Moreover, the use of Deny and Counter in English is clearly separated, but Spanish authors use both categories almost equally. Consequently, English authors have a major preference for contrastive expressions such as *but*, implying that they will introduce aspects positive and negative to usually soften their comments, while Spanish authors provide stronger affirmative and negative (hence the higher use of Deny) statements.

With respect to Proclaim, it is, together with the Focus category in Graduation, as will be seen later, the least used Appraisal tag in this type of text. Their results do not



exceed 5%; more specifically, Concur obtained 2.45% and 0.89%, Pronounce, 3.68% vs. 3.35%, and Endorse, 0.77% and 0.00% in English and Spanish, respectively. Due to the short type of text this genre is, authors do not typically include allusions to other people's opinion, no matter whether they agree or disagree with them. It must be remembered that the rating system in Play Store displays a list of all the comments already introduced in the system. Consequently, authors do not find it necessary to mention other people's opinions, since any user can read hundreds of other people's comments themselves. Instead, review authors most often focus on their own opinion, either more or less strongly by means of Entertain resources, and item's qualities.

The most frequent Engagement spans are listed in Table 19.

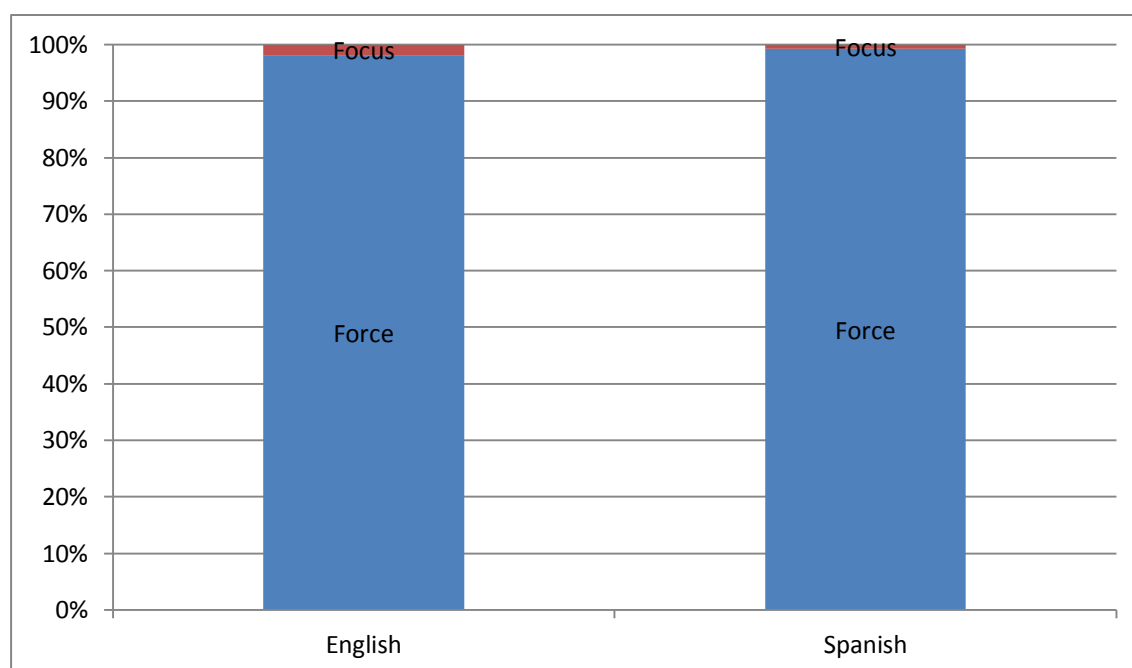
English			Spanish		
#	%	Span	#	%	Span
85	13.04	but	127	28.35	no [no, not]
54	8.28	not	59	13.17	pero [but]
34	5.21	just	25	5.58	si [if]
33	5.06	if	14	3.12	solo [only]
19	2.91	only	9	2.01	nada [nothing]
16	2.45	don't	9	2.01	único [only]
15	2.30	even	8	1.79	parece [seems]
14	2.15	didn't	7	1.56	sin [without]
14	2.15	still	7	1.56	ni [and-no]
13	1.99	could	6	1.34	simplemente [just]

Table 19. Most common Engagement spans in English and Spanish

Most of these spans belong to the categories Deny and Counter, as could be expected. Shared expressions in Counter are exemplified by *but&pero*, *just&simplemente*, *if&si*, *only&solo/único*, and English includes *could*, *still* and *even* additionally. The use of Deny presents a bit more complex range of expressions since the particle used for negation in English is merged with the auxiliary form. Thus, *not*, *don't* and *didn't* are used in negative sentences, the same as the adverb *no* [no, not] in Spanish. This language does also include other negative items: *ni* [and-no], *nada* [nothing], and *sin* [without]. There is only one single span that does not belong to Counter or Deny: *parece* [seems]. This span was classified as Evidential. As said above, the use of Evidential markers was much more common in Spanish than in English, which was richer in Epistemic spans.

### 4.4.3. Graduation

Graduation is the third and last of the main three axes of Appraisal Theory. Resources to intensify or soften other words such as nouns, adjectives or verbs and the concepts they represent are included here. The distribution of the two types of Graduation is shown in Figure 28.



**Figure 28. Distribution of Graduation in English and Spanish**

Force does not only represent the third most used category of Appraisal, behind Contraction and Appreciation, but it also covers most of the Graduation type. The realisations of Focus are less than 2% of Graduation in both languages. However, not all the minor categories in Force are evenly distributed, as shown in Figure 29.

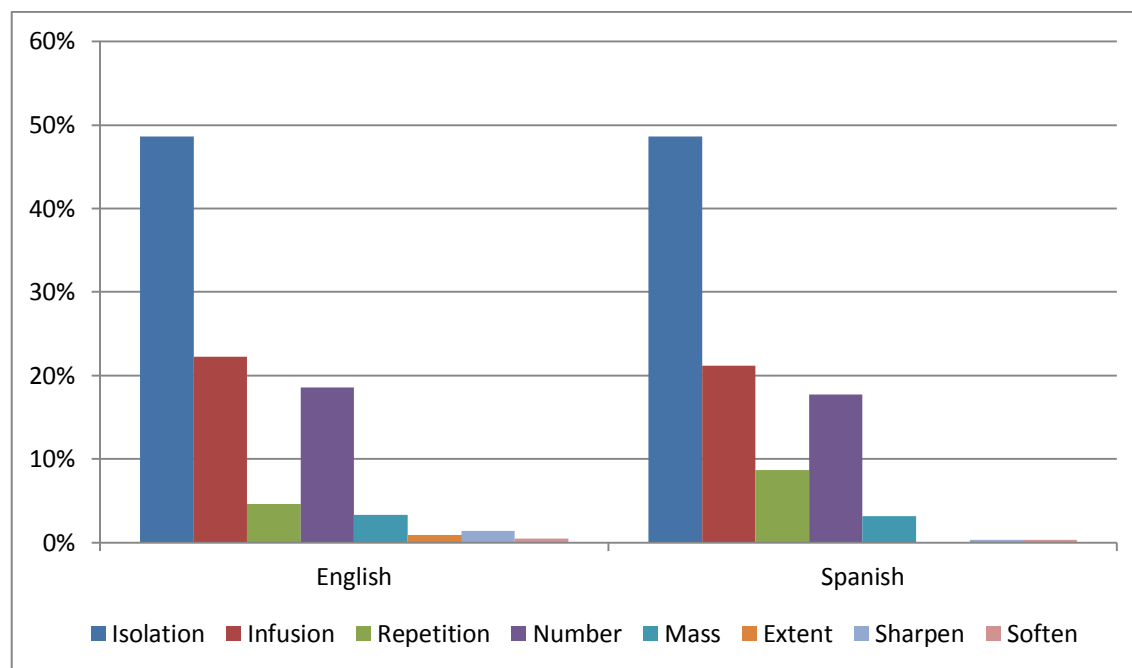


Figure 29. Distribution of Force and Focus in English and Spanish

There are no significant differences in the use of Graduation in English and Spanish, except for the use of Repetition, which is relatively higher in Spanish (4.59% vs. 8.68%). This is basically due to the use of a generous number of question and exclamation marks in the Spanish texts. It must be explained that, unlike English, the Spanish language uses question and exclamation marks not only to close their sentences but also to open them. However, only closing marks were annotated in the texts, in order to obtain comparable results.

The results for the rest of categories are as follows: Isolation (48.62% vs. 48.61%), Infusion (22.25% vs. 21.18%), Number (18.58% vs. 17.71%), Mass (3.21% vs. 3.12%) and Extent (0.92% vs. 0.00%). As can be seen, Isolation resources (*so*, *very*, *totally*, *súper* [super], *bastante* [quite], *un montón* [a lot], etc.) cover most of the Force type. Infusion is the second most used type, with spans such as *every time*, *greatest*, *sometimes*, *mejor* [better, best], *siempre* [always], *peor* [worse, worst], etc.). It is important to mention that Infusion does also include capital letters when used to highlight or emphasise a particular span since these do not imply a Force meaning unless they contrast to small letters intentionally. Number makes use of quantifiers such as *many*, *few*, *plenty of*, *several*, *muchos* [many], *medio* [half], etc. Some Mass examples include *big*,

*large*, *pequeño* [little], *enormemente* [hugely]. Finally, the few Extent realisations are illustrated by *wide*, *short* and *narrow*.

Regarding the Focus categories, Sharpen (1.38% vs. 0.35%) and Soften (0.46% vs. 0.35%), hardly any occurrences were annotated, although they were more common in English, which concurs with the higher results in Epistemic spans obtained in English texts. Focus realisations include *pure*, *kinda*, *genuinely* and *mas o menos* [more or less].

Table 20 lists the most common spans in Graduation.

English			Spanish		
#	%	Span	#	%	Span
35	8.03	!	42	14.58	muy
25	5.73	so	21	7.29	!!
20	4.59	more	16	5.56	mejor
19	4.36	most	16	5.56	más
18	4.13	very	14	4.86	mas
13	2.98	much	12	4.17	!
13	2.98	really	10	4.17	mucho
13	2.98	too	8	3.47	siempre
12	2.75	many	6	2.78	tan
11	2.52	some	6	2.08	!!!

Table 20. Most common Graduation spans in English and Spanish

Highly noticeable is the use of exclamation marks in the Spanish texts. Spanish authors do not only use single exclamation marks as their more frequent Graduation resource but also double and even triple marks. English authors do also use repeated exclamation marks but in a visibly lower proportion. This supports previous findings regarding the strength of Spanish comments in contrast to English texts. They do not only make stronger statements but they also introduce explicit emphasis by means of exclamation marks.

Other shared resources include *so&tan*, *more / most&más (mas)*, *very&muy* and *much&mucho*. *Really*, *too*, *many* and *some* are also very frequent in English, while the Spanish texts show the comparative and superlative *mejor* more often than their counterparts better and best, as well as *siempre* is more recurrent than *always*.

### 4.5.Item differences (applications vs. games vs. books vs. films vs. music)

Some remarkable differences have been pointed out in the use of Appraisal in English and Spanish. However, the texts addressed several products, and these products may involve Appraisal differences in the way authors review them. The products analysed were applications, games, books, films and music. In order to compare Appraisal occurrences in each of these categories, five groups of 40 texts each (20 in English and 20 in Spanish) were made, and statistics for the marking of spans in each of them were extracted. Figure 30 presents the results for Appraisal annotation for these five products.

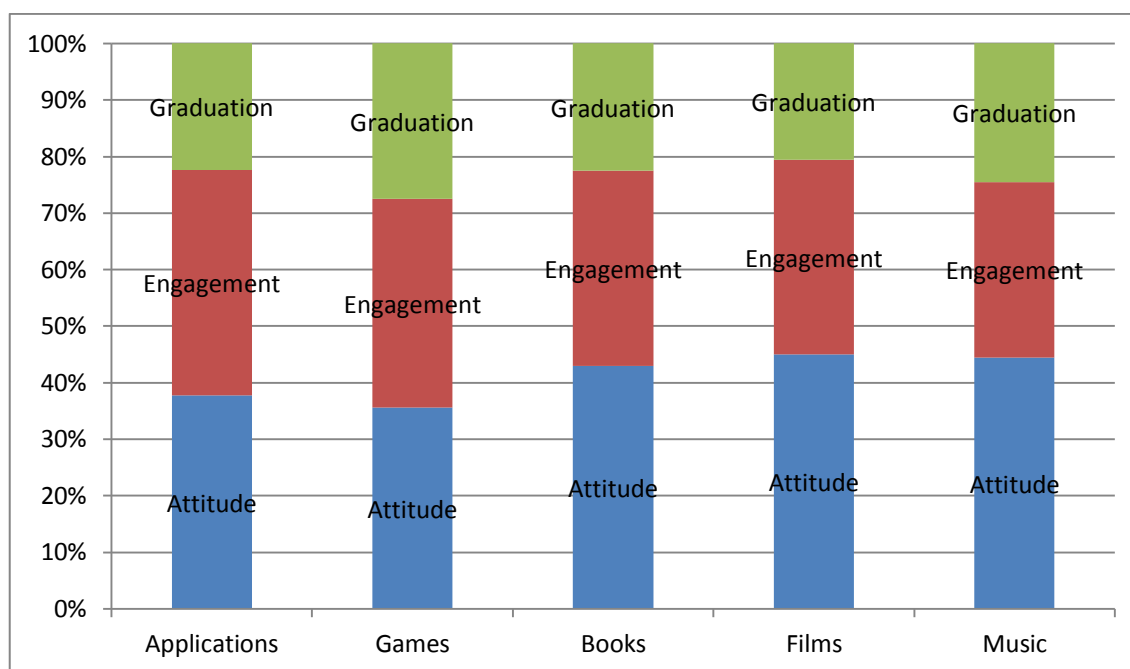


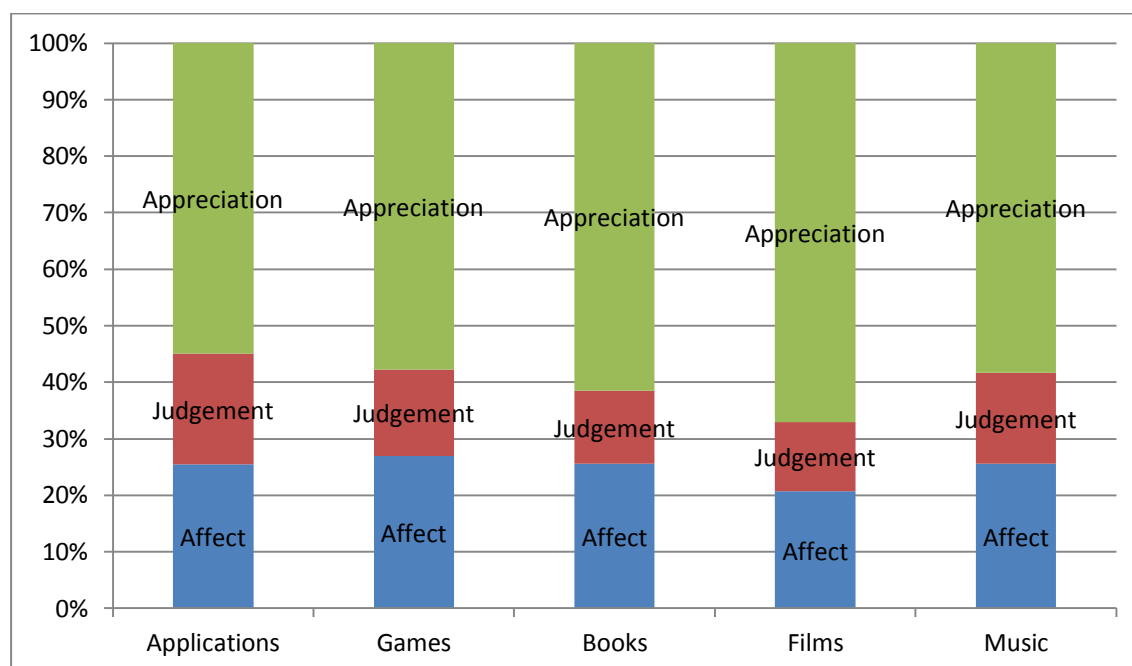
Figure 30. Distribution of Appraisal in applications, games, books, films and music

A quick glance at these initial results shows a small difference between two groups: applications and games on the one hand, and books, films and music on the other. The group of applications and games characterised by a higher use of Engagement resources, while the group of books, films and music shows a more frequent use of Attitude categories. Graduation does not show a clear tendency and presents similar results except for a wider use in games.

The following sections will analyse these results in more detail including more specific Appraisal categories.

### 4.5.1. Attitude

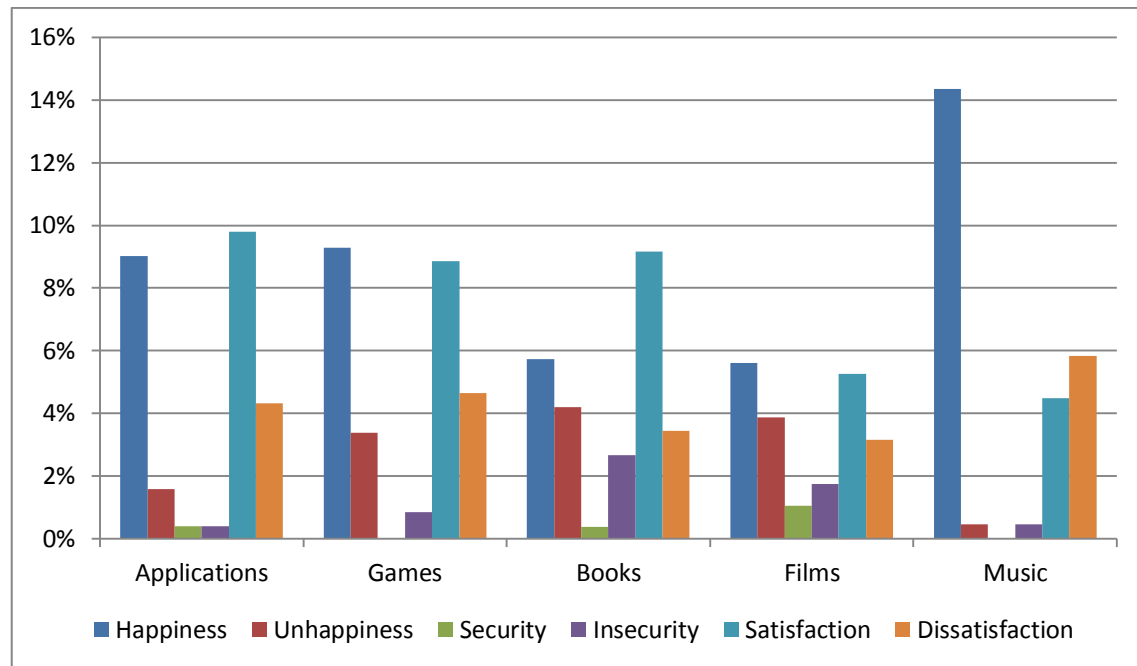
Attitude is the most repetitive resource in the five types of texts, although the inner distribution of the tagged spans may present some variations. These results are displayed in Figure 31.



**Figure 31. Distribution of Attitude in applications, games, books, films and music**

Appreciation is ranked first as an Attitude resource, although some reviewed items are more prone to include them than others. This is the case of films, in which the proportion of Appreciation is the highest out of the five items, followed by books. Affect is realised more frequently in games, as well as in applications and books. And Judgement is more likely to appear in applications than in any other item reviewed.

Each minor category is analysed below, starting with Affect in Figure 32.



**Figure 32. Distribution of Affect in applications, games, books, films and music**

Regarding Happiness, music has obtained clear exceptional results (14.80%); a second group formed by application and games has average Happiness results (9.02% vs. 9.28%); a third group, books and films, has the lowest Happiness spans (5.73% vs. 5.61%). The second most used Affect category is Satisfaction, and is most present in applications (9.80%), games (8.86%) and books (9.16%), while films and music obtained lower results (5.26% vs. 4.48%). This may be related to the type of feeling that each product is usually linked to: authors pay more attention to how satisfied they are with applications and games—usually regarding their performance—and books and films—usually with respect to the plot, characters, etc., in a lower degree—, while music transmits a different kind of feelings, namely Happiness-related.

The results for Dissatisfaction are, nevertheless, more evenly distributed: 4.31% (applications), 4.64% (games), 3.44% (books), 3.16% (films) and 5.83% (music). Together with Dissatisfaction, Unhappiness is also present in most of the products (1.57% vs. 3.38% vs. 4.20% vs. 3.86% vs. 0.45%). The lower results of Unhappiness in music in combination with the high results in Dissatisfaction show that when authors want to support an artist they refer to positive Happiness feelings, while when criticising negatively, they address Dissatisfaction issues. In contrast, in the rest of products, both Happiness and Satisfaction are higher than their negative counterparts.

Security and Insecurity obtained, as seen before, the lowest results as Affect categories. This type of feelings does not typically occur in this genre, no matter the type of product reviewed, especially Security: 0.39% (applications), 0.00% (games), 0.38% (books), 1.05% (films) and 0.00% (music). The only exceptions found are Books (2.67%) and films (1.75%), where the Insecurity rate is slightly higher than in applications (0.39%), games (0.84%) and music (0.45%).

Generally speaking, book and films are poorer than the other categories in positive aspects such as Happiness or Satisfaction, while they are richer in negative resources such as Unhappiness and Insecurity. This may be interpreted as a higher interest in criticising negatively that kind of products, while authors spend more time on complimenting applications, games and music.

The second major category in Attitude is Judgement, whose results are presented in Figure 33.

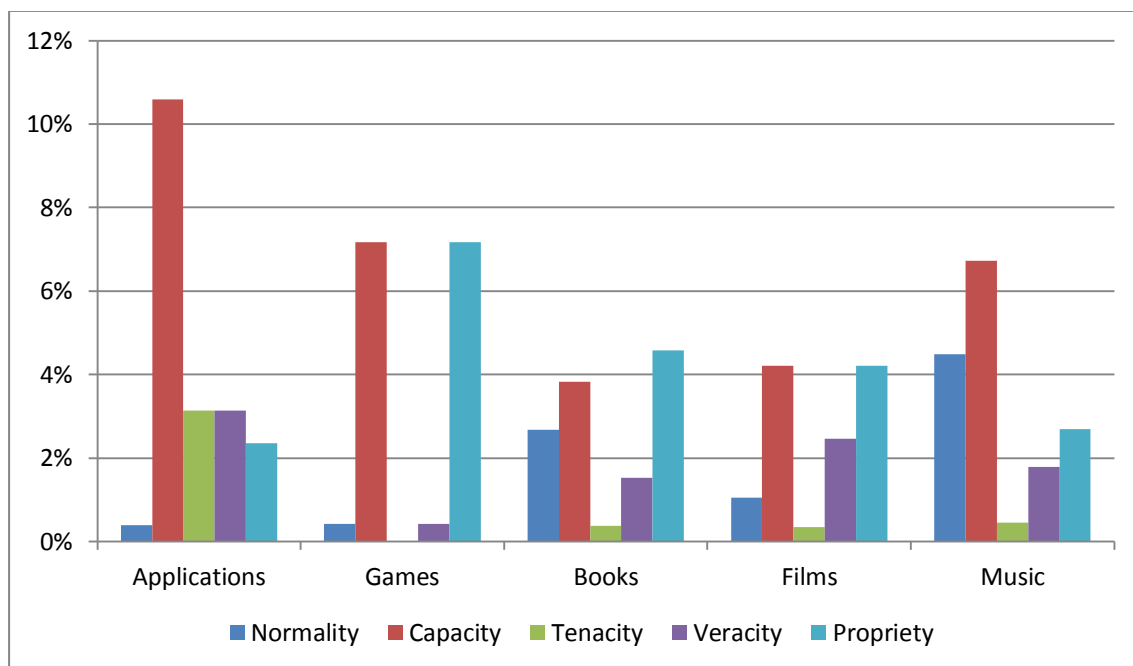


Figure 33. Distribution of Judgement in applications, games, books, films and music

Applications present a higher interest in Capacity aspects (10.59%), dealing with information about requirements and performance. Thus, Capacity is also an important Judgement resource for Games (7.17%) and Music (6.73%), since the performance and talent of the specific game or artist is a core issue. Books (3.82%) and Films (4.21%), on



the other hand, do not pay so much attention to this kind of feature where, in general, the results for Judgement are lower.

As has been already mentioned, Propriety usually includes negative meanings regarding moral aspects. It is surprising to find them when assessing objects, like those included in the study, but users do actually use them. More specifically, Games have obtained high results in Propriety (7.17%) because authors complain about bugs that occur in the game and make users lose their bonuses or even about the number or price of the purchases included in the game. The results for Propriety in the rest of categories are lower: 2.35% (Applications), 4.58% (Books), 4.21% (Films) and 2.69% (Music).

Tenacity and Veracity meanings obtained a very modest representation: 3.14% vs. 3.14% (Applications), 0.00% vs. 0.42% (Games), 0.38% vs. 1.53% (Books), 0.35% vs. 2.46% (Films) and 0.45% vs. 1.79% (Music). These are not qualities that users tend to appreciate in these types of products.

Normality results are higher in Music (4.48%) because authors usually mention the value of the artist as someone different (or similar) to other artists or previous compositions by the same artist. Since this information is not provided in the rest of products, Normality results are lower: 0.39% (Applications), 0.42% (Games), 2.67% (Books) and 1.05% (Films).

Finally, Appreciation subtypes are covered in Figure 34.

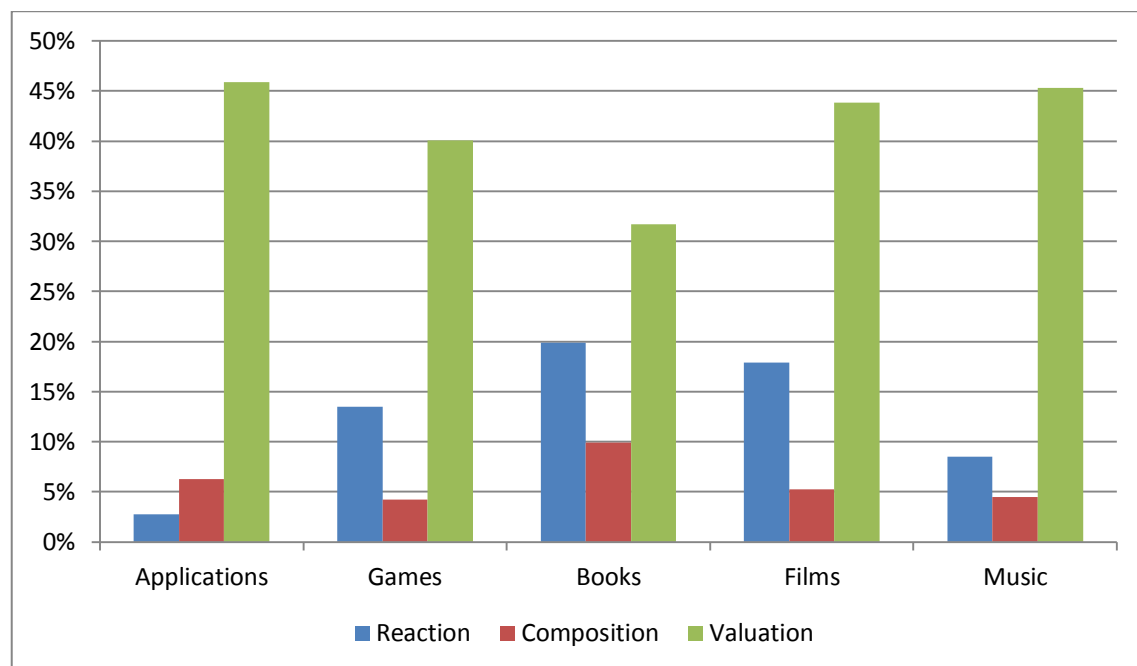


Figure 34. Distribution of Appreciation in applications, games, books, films and music

Valuation has clearly obtained the widest results in Appreciation, followed by Reaction. What is more, a visible tendency is observed in the inverse proportion of these two types of resources: the higher the Valuation results, the lower the Reaction ones. The specific percentages for Reaction and Valuation are the following ones: 2.75% vs. 45.88% (Applications), 13.50% vs. 40.08% (Games), 19.85% vs. 31.68% (Books), 17.89% vs. 43.86% (Films) and 8.52% vs. 45.29% (Music). This means that the more the author focuses on the particular qualities of the item reviewed, the less they care about what the impact it has on users, and vice versa. Authors tend to focus on Valuation more usually, but in case they do not, they include Reaction expressions to complement and describe their opinions.

As expected, Composition obtained the lowest results in Appreciation, unable to exceed 10% in any of the products: 6.27% (Applications), 4.22% (Games), 9.92% (Books), 5.26% (Films) and 4.48% (Music).

#### 4.5.2. Engagement

Engagement statistics arranged by type of product are presented in Figure 36.

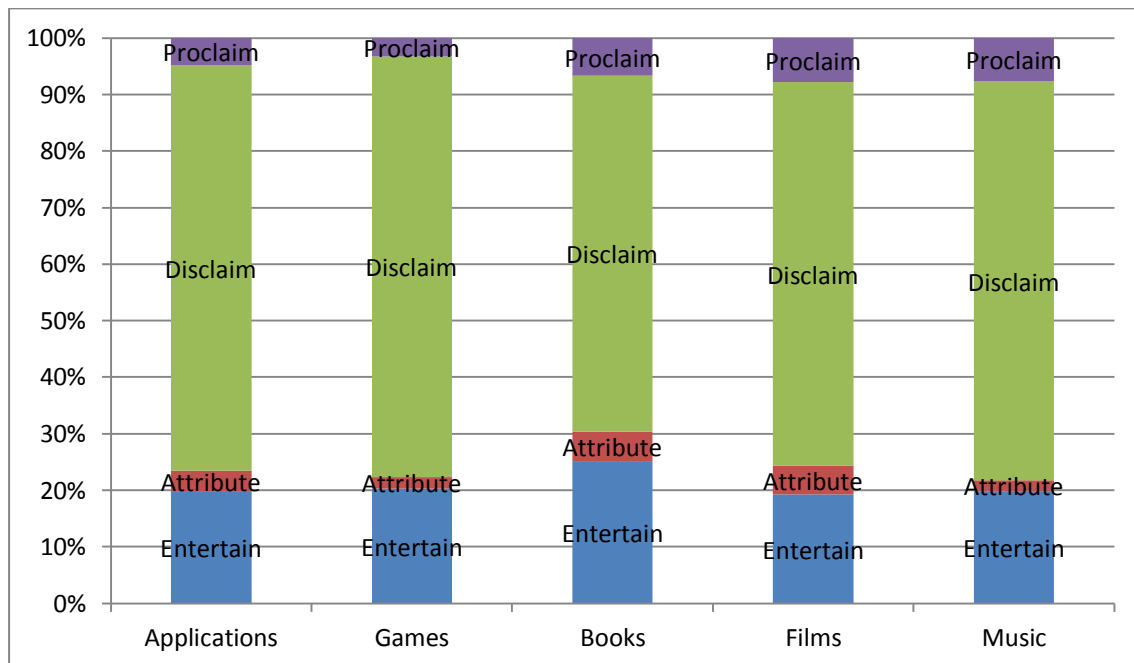
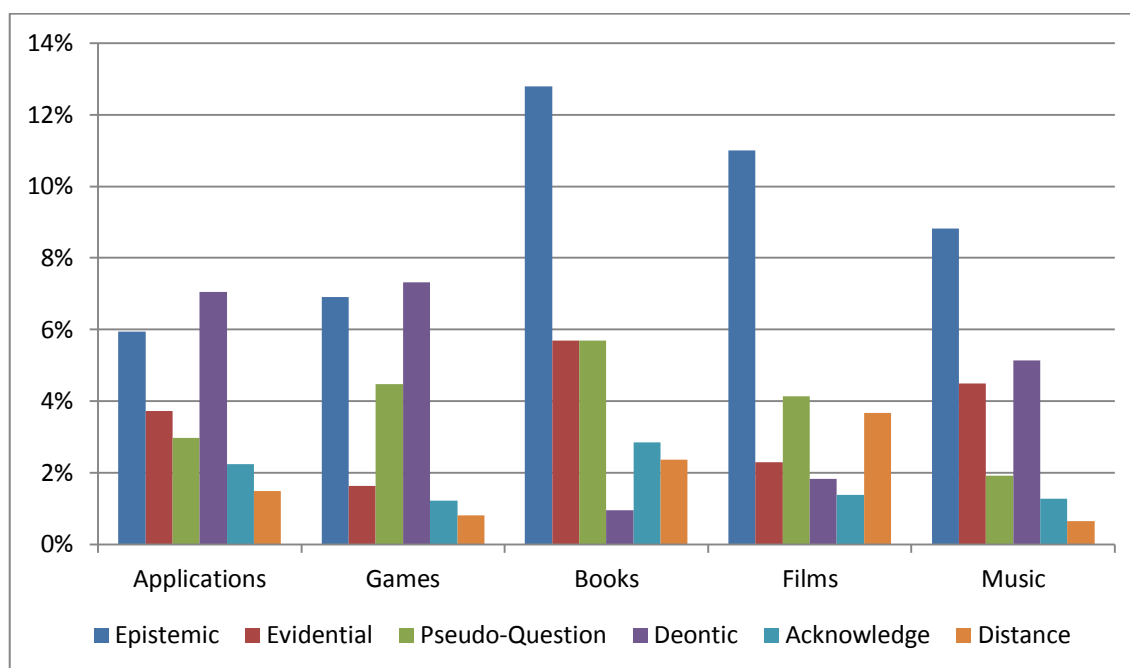


Figure 35. Distribution of Engagement in applications, games, books, films and music

There are two tendencies in Engagement results. On the one hand, Disclaim resources are more abundant in applications and games. On the other hand, Proclaim expressions are more typically found in books, films and music. The results for Expansion types (Entertain and Attribute) are similar among the five products to some extent, with the exception books where there may be a higher interesting in this specific type. Further information on this axis will shed some light on these tendencies. Firstly, the results for Expansion types are presented in Figure 37.



**Figure 36 Distribution of Expansion in applications, games, books, films and music**

The higher use of Entertain in books observed in the previous figure is due to the visible different results obtained in Epistemic resources (12.80%), which is not so high in the rest of products: 5.95% (applications), 6.91% (games) and 8.33% (music). Epistemic spans are also frequent in films (11.01%), but the general results were not so high because the rest of Expansion categories are lower.

As a tendency, Epistemic and Deontic resources seem to be related: the higher Epistemic results, the lower the Deontic. The Deontic results are 7.06% (applications), 7.32% (games), 0.95% (books), 1.83% (films) and 5.13% (music). Applications and games show similar results for both categories, with a slight preference for Deontic spans, while books, films and music do clearly lean on Epistemic in detrimental to Deontic ones.

The rest of features present irregular results among the products. The two Entertain types pending, Evidential and Pseudo-Questions, obtained the following results: 3.72% vs. 2.97% (applications), 1.63% vs. 4.47% (games), 5.69% vs. 5.69% (books), 2.29% vs. 4.13% (films) and 4.49% vs. 1.92% (music). Applications and music show higher results in Evidential markers while games and films prefer Pseudo-Question items. In addition, Books shows similar results for both categories. The percentages of spans annotated as Attribute categories, that is, Acknowledge and Distance, are as follows: 2.23% vs. 1.49% (applications), 1.22% vs. 0.81% (games), 2.84% vs. 2.37% (books), 1.38% vs. 3.67% (films) and 1.28% vs. 0.64% (music). They all obtained a representation lower than 4% and most of them lower than 3%. Consequently, no significant differences can be pointed out regarding the use of Attribute according to the product reviewed.

Secondly, Contraction types are illustrated in Figure 38.

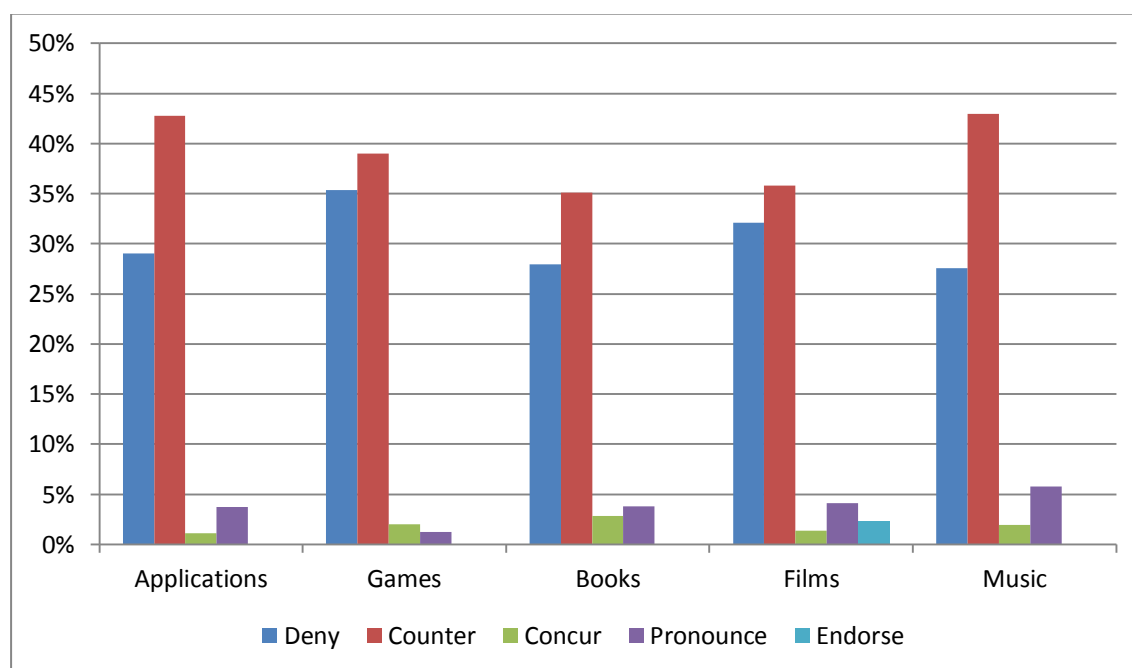


Figure 37. Distribution of Contraction in applications, games, books, films and music

All the products show a preference for Counter resources than for Deny elements. However, applications and music show similar results for both Deny and Counter categories: 29.00% vs. 42.75% in applications, and 27.56% vs. 42.95% in the case of music. On the other hand, games, books and films have obtained different proportions in their results: 35.37% vs. 39.02% (games), 27.96% vs. 35.07% (books) and 32.11% vs.

35.78% (films). No pattern is followed apart from the specific requirements of each product.

As happened with Attribute types on Expansion, Proclaim types in Contraction show exceptionally low results compared to Disclaim categories, and no tendencies are manifested in the spans tagged as Proclaim. The results for Concur, Pronounce and Endorse are the following ones: 1.12% vs. 3.72% vs. 0.00% (applications), 2.03% vs. 1.22% vs. 0.00% (games), 2.84% vs. 3.79% vs. 0.00% (books), 1.38% vs. 4.13% vs. 2.29% (films) and 1.92% vs. 5.77% vs. 0.00% (music). Generally speaking, it can be said that there is a slight preference for Pronounce expressions over Concur (in all products except for games), and Endorse is the least represented category with results equal to zero in three of the five products.

Consequently, no interest is observed in the texts' authors to incorporate authorial external voices or provide this authority to their own voices. Authors reflect their opinions on their texts without so much thinking of proving them or introducing undeniable statements.

### 4.5.3. Graduation

The results for the third major Appraisal axis, Graduation, are presented in Figure 39.

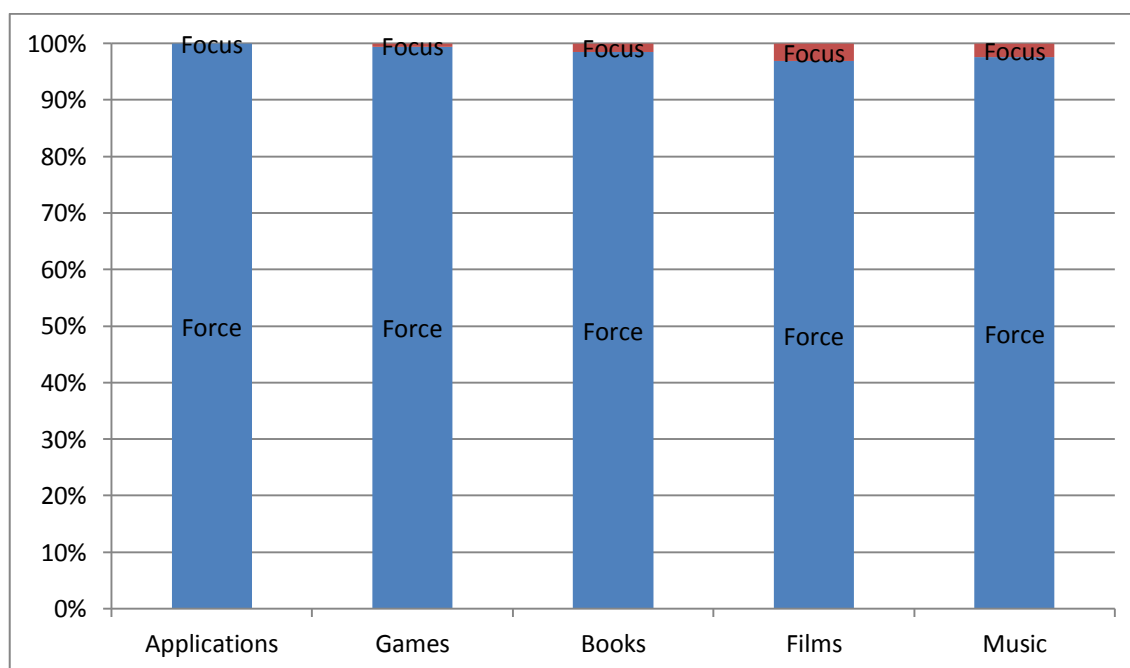
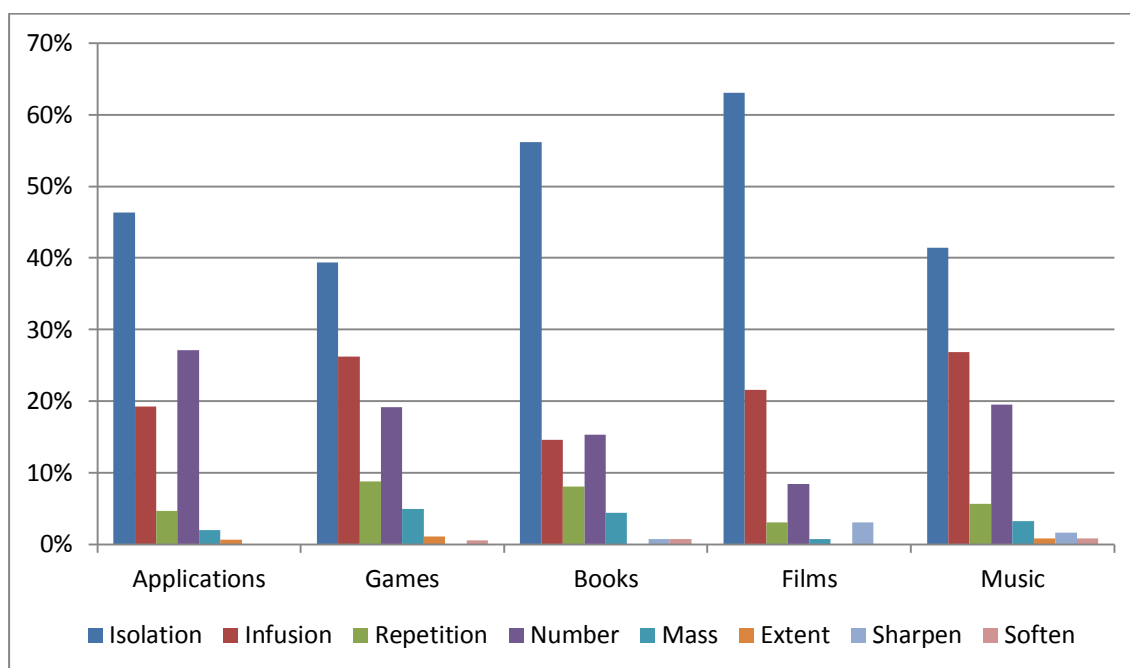


Figure 38. Distribution of Graduation in applications, games, books, films and music

Although some differences can be seen among the five items, they are not significant enough due to the low total number of Focus occurrences (0.00% vs. 0.55% vs. 1.46% vs. 3.08% vs. 2.44%). They all show how Force is far more common than Focus in this genre, independently from the item reviewed (100% vs. 99.45% vs. 98.54% vs. 96.92% vs. 97.56%).

Figure 40 shows the distribution of Graduation minor categories, namely, Force and the scarce Focus.



**Figure 39. Distribution of Force and Focus in applications, games, books, films and music**

Regarding minor categories in Graduation, some variation is shown in the use of Isolation resources. Books and films are richer in this category (56.20% vs. 63.08%) than the other three categories (Applications 46.36% vs. Games 39.34% vs. Music 41.46%), while games and music make a wider use of Infusion terms (26.23% vs. 26.83%) than applications, books and films (19.21% vs. 14.60% vs. 21.54%).

At the same time, books and films are poor in Number elements (15.33% vs. 8.46%), while applications, games and music incorporate them to their lexis more often (27.15% vs. 19.13% vs. 19.51%). Repetition resources are more frequently used in Games and Books (8.74% vs. 8.03%) than in applications, films and music (4.64% vs. vs. 3.08 vs. 5.69%). Similar results were obtained for Mass spans, which add up to 4.92% in

games and 4.38% in books, while only 1.99%, 0.77% and 3.25% in applications, films and music, respectively.

Finally, the last Force category is Extent, whose results are close to or actually zero: 0.66% (applications), 1.09% (games), 0.00% (books), 0.00% (films) and 0.81% (music). A similar situation is found for Focus categories, in which Sharpen and Soften had 0.00% vs. 0.00% (applications), 0.00% vs. 0.55% (games), 0.73% vs. 0.73% (books), 3.08% vs. 0.00% (films) and 1.63% vs. 0.81% (music).

To sum up, books and films pay more attention to Isolation in detrimental to Number, that is, authors include resources such as the words *very* and *so* other elements like exclamation marks in their comments, but not so many quantifiers such as *many* or *some*. Games and music, even though their most frequent resource is still Isolation, have shown higher values for Infusion than the rest of categories, this pointing to a wider interest in comparative and superlative adjectives, capital letters, and other Infusion resources. Games and books are more profuse in Repetition (use of a single word or element several times (e.g. *way way way*), or letter lengthening (*soo*), and Mass spans (*big, huge, small*, etc.). Finally, Extent, Sharpen and Soften have hardly been represented, as explained in previous sections. Therefore, the results are very distributed among the different categories and no clear tendencies can be pointed apart from individual preferences in each of the items reviewed.

## 4.6. Polarity and rating differences

Reviews can also be classified according to the number of stars they assign to the item reviewed. This rating goes from 1 to 5 stars, being 5 the best valuation. This study analysed comments with 1, 2, 4 and 5 stars. In this section, the texts are organised in two groups: one includes 100 texts with 1 and 2 stars (negative reviews) while the other includes reviews 100 texts with 4 and 5 stars (positive reviews).

The distribution of Appraisal tags in the texts arranged according to this criterion is presented in Figure 41.

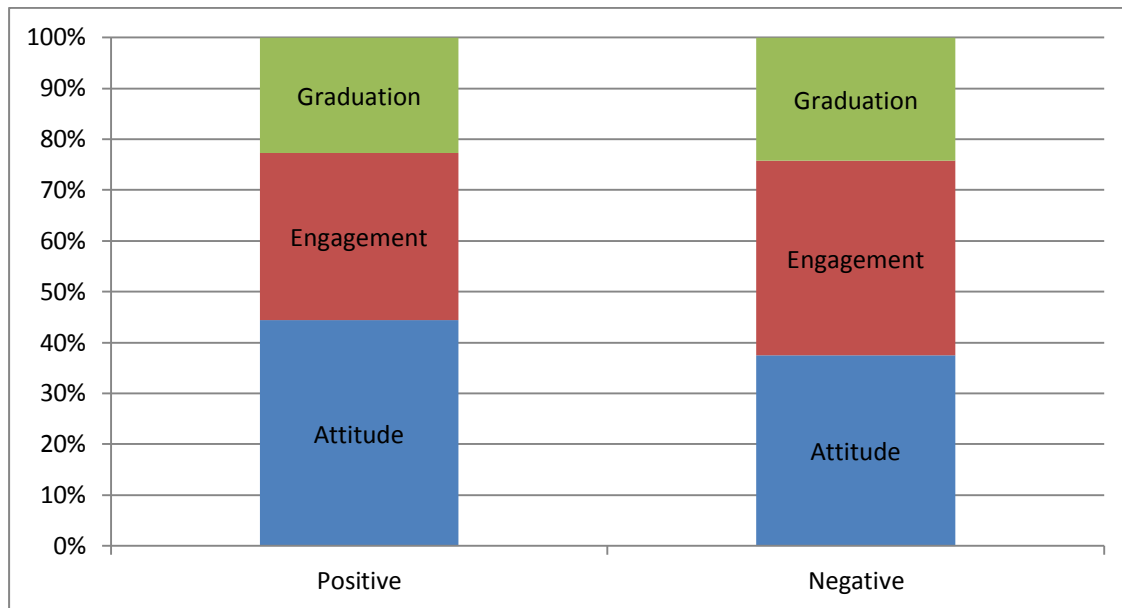


Figure 40. Distribution of Appraisal in positive and negative reviews

A slight difference in the use of Attitude and Engagement can be perceived: while Attitude resources are more abundant in positive reviews, negative reviews use Engagement expressions more than their counterpart. The use of Graduation, on the other hand, is virtually the same in both types of texts.

#### 4.6.1. Attitude

Global Attitude results are presented in Figure 42.

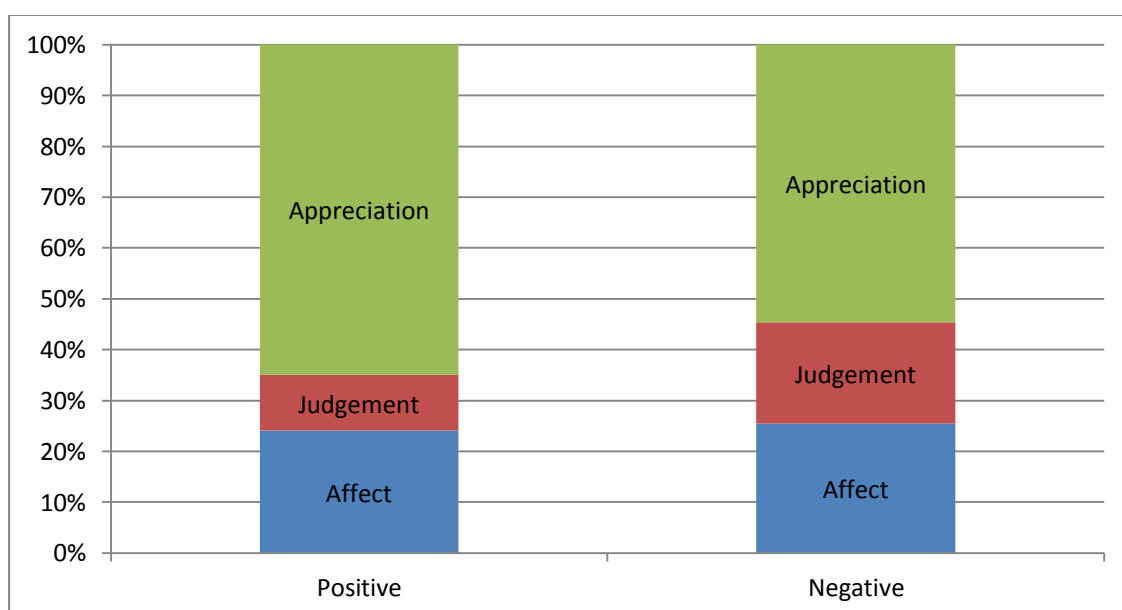


Figure 41. Distribution of Attitude in positive and negative reviews



Although the distribution of Affect is almost identical in positive and negative reviews (24.07% vs. 25.47%), Appreciation (64.93% vs. 54.67%) and Judgement (11.00% vs. 19.86%) show visible differences. While positive reviews are more profuse on the former, negative reviews make a wider use of Judgement types. This is related to the specific use that authors make of Judgement expressions, focusing on Propriety and, more specifically, negative spans, like those presented in section 4.4.1. in relation to Figure 23. In addition, when users are giving a positive evaluation of the item, they use more Appreciation resources to describe and justify the positive rating.

Affect results are presented in Figure 43.

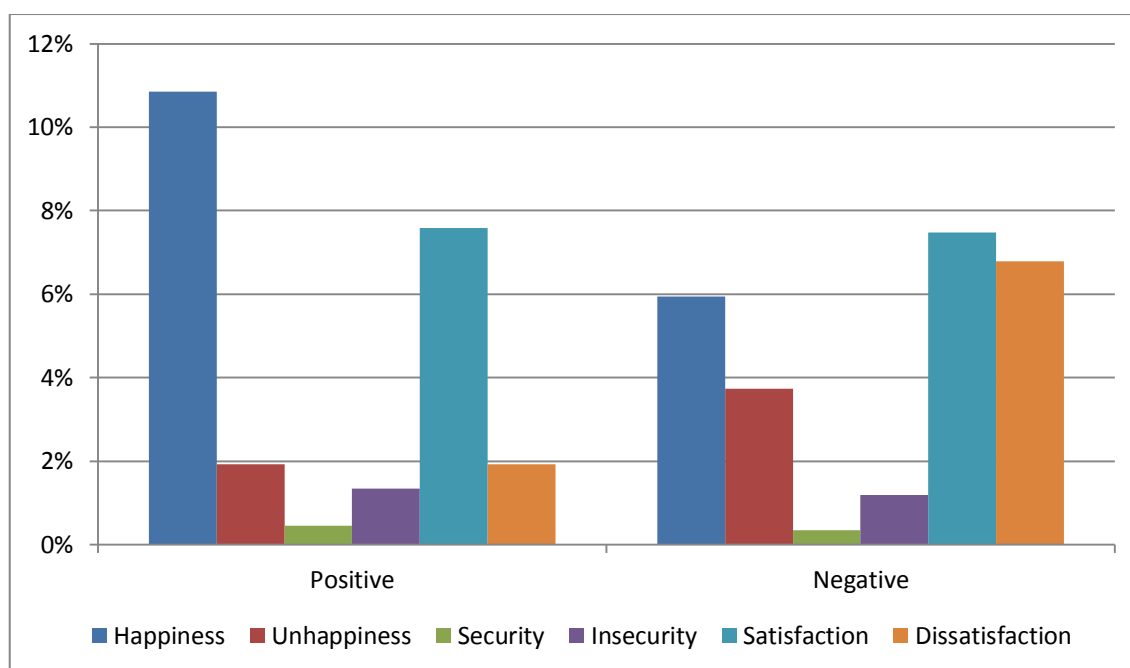


Figure 42. Distribution of Affect in positive and negative reviews

Positive and negative reviews show a clear difference in the use of Happiness (10.85% vs. 5.94%) vs. and Dissatisfaction (1.93% vs. 6.79%) resources: when criticising, authors will be more prone to include Dissatisfaction expressions such as *complaint*, *disappointed* or *aburrir* [bore], but when complimenting, it is not Satisfaction that authors use, but Happiness spans such as *love*, *amo* [(I) love], *gusta* [like], *encanta* [love], *conmovió* [(he/she/it) moved]. In the same way, Unhappiness annotations like *hated*, *dislike* or *miss*, appear more frequently in negative than positive reviews but, surprisingly, Satisfaction manifests evenly in both types of texts, including spans like *enjoy*, *interested* or

*recomiendo* [(I) recommend]. More marginal categories include Security (0.45% vs. 0.34%) and Insecurity (1.34% vs. 1.19%), as has been observed in previous analyses.

Judgement results are presented in Figure 44.

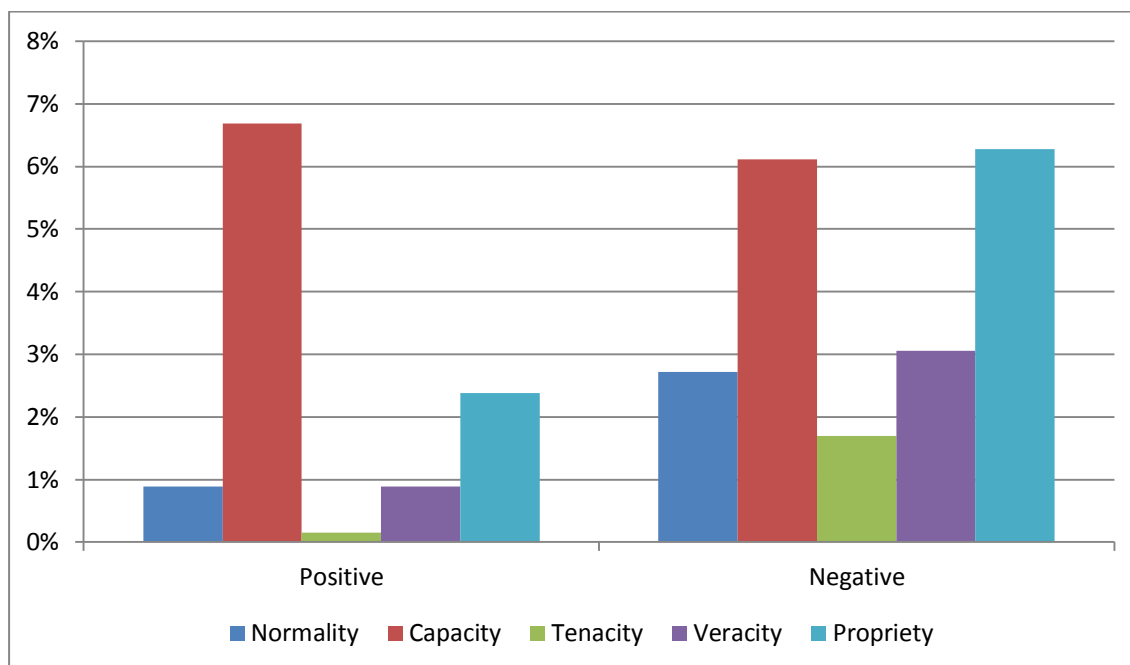
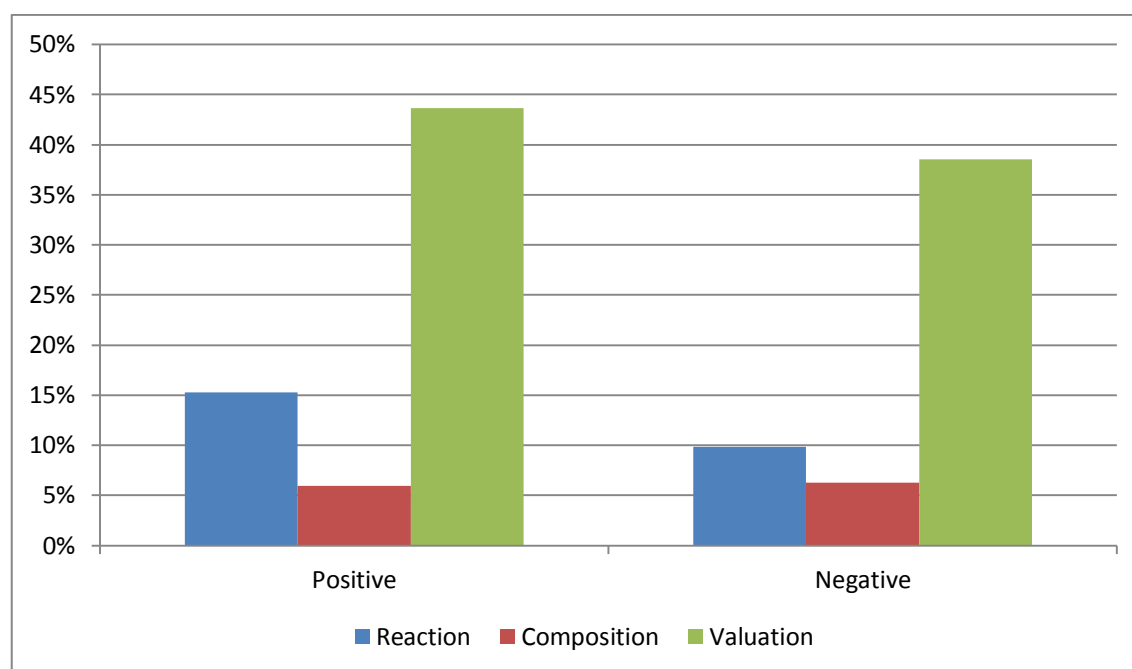


Figure 43. Distribution of Judgement in positive and negative reviews

As mentioned above, most of the more specific categories in Judgement are more abundant in negative comments than in positive ones. The only exception is Capacity (6.69% vs. 6.11%), which is used both in negative and positive reviews. In addition, the use of Propriety in texts giving a negative rating is visibly higher than in their positive counterparts (2.38% vs. 6.28%). Propriety is a category from which mainly negative expressions are employed in reviews. This may be due to the fact that moral qualities are typically taken for granted, that is, users will not usually admire how moral someone is because they are expected to behave morally; however, when someone deviates from the established rules, then negative Propriety resources are used. The percentages for the rest of categories are as follows: Normality (0.89% vs. 2.72%), Tenacity (0.15% vs. 1.70%) and Veracity (0.89% vs. 3.06%).

The last major subsection of Attitude, namely, Appreciation, is presented in Figure 45.

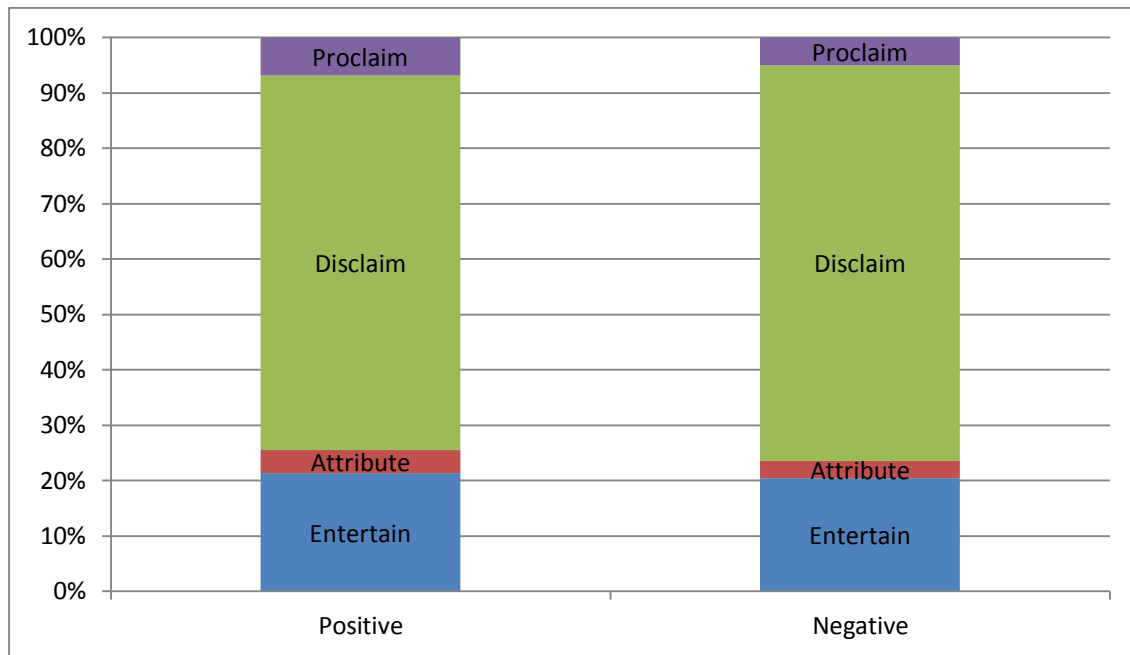


**Figure 44. Distribution of Appreciation in positive and negative reviews**

The proportion among the three categories is similar in both positive and negative texts. However, more examples are found in the former than in the latter: Reaction (15.30% vs. 9.58%), Composition (5.94% vs. 6.28%) and Valuation (43.68% vs. 38.54). This is probably due to the higher results obtained in Judgement in negative reviews. These results point to a heavier weight of moral aspects to reject the items reviewed but a preference for other values to rate them positively.

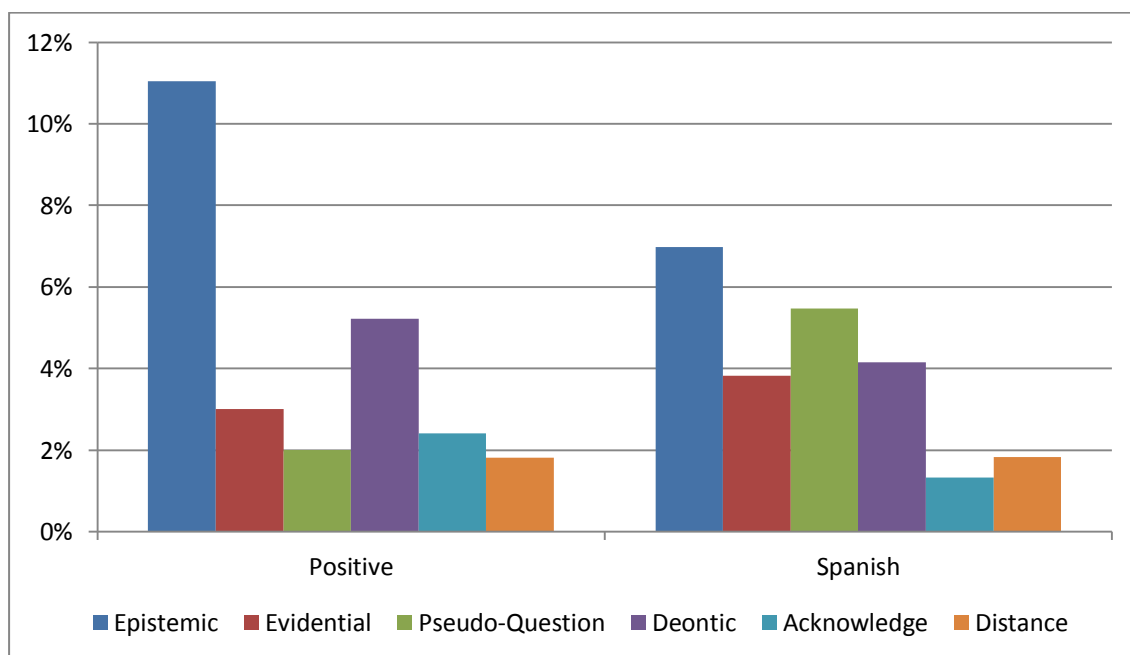
### 4.6.2. Engagement

General results for the second major axis of Appraisal, Engagement, are presented in Figure 46.



**Figure 45. Distribution of Engagement in positive and negative reviews**

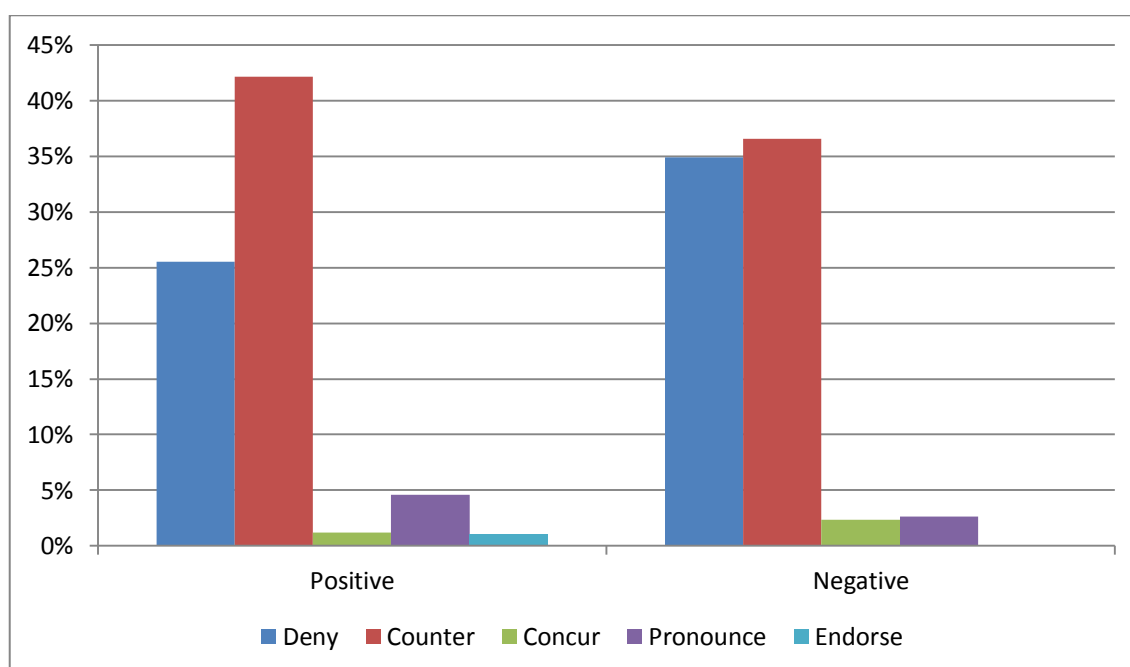
No apparent difference is observed between positive and negative reviews regarding the use of Engagement features: Entertain (21.29% vs. 20.43%), Attribute (4.22% vs. 3.16%), Disclaim (67.67% vs. 71.43%) and Proclaim (6.83% vs. 4.98%). However, a closer look at the minor categories inside this axis will show some visible distinctions. Firstly, Figure 47 summarises the results of Expansion types.



**Figure 46. Distribution of Expansion in positive and negative reviews**

The use of Epistemic resources is evidently higher in positive reviews (11.04% vs. 6.98%). This means that authors address the certainty of their statements—either to indicate a higher or a lower level of certainty— more frequently when talking positively than in the opposite situation. On the other hand, the use of Pseudo-Questions is clearly linked to negative comments (2.01% vs. 5.48%), which express authors’ feelings or surprise and rejection of a particular feature or aspects related to the item reviewed, as explained in previous results. The results for other categories do not show significant differences: Evidential (3.01% vs. 3.82%), Deontic (5.22% vs. 4.15%), Acknowledge (2.41% vs. 1.33%) and Distance (1.81% vs. 1.83%).

Secondly, Figure 48 displays Contraction results.



**Figure 47. Distribution of Contraction in positive and negative reviews**

Proclaim resources are almost inexistent in both types of texts. This category includes Concur (1.20% vs. 2.33%), Pronounce (4.62% vs. 2.66%) and Endorse (1.00% vs. 0.00%). However, evident differences are found in Disclaim elements. While positive comments lean on Counter expressions more frequently, negative reviews rely on Deny items almost as much as on Counter ones: Deny and Counter results in positive results add up to 25.50% and 42.17% respectively, and 34.88% and 36.54% in negative reviews. The use of Deny items is strictly related with negation, which is expected to play an

important role in texts with a negative rating. Moreover, these texts do not allow for hypothetical situations so much, but they focus on what actually happens, lowering the number of Counter occurrences.

### 4.6.3. Graduation

The results of the annotation of Graduation features in positive and negative reviews are presented in Figure 49.

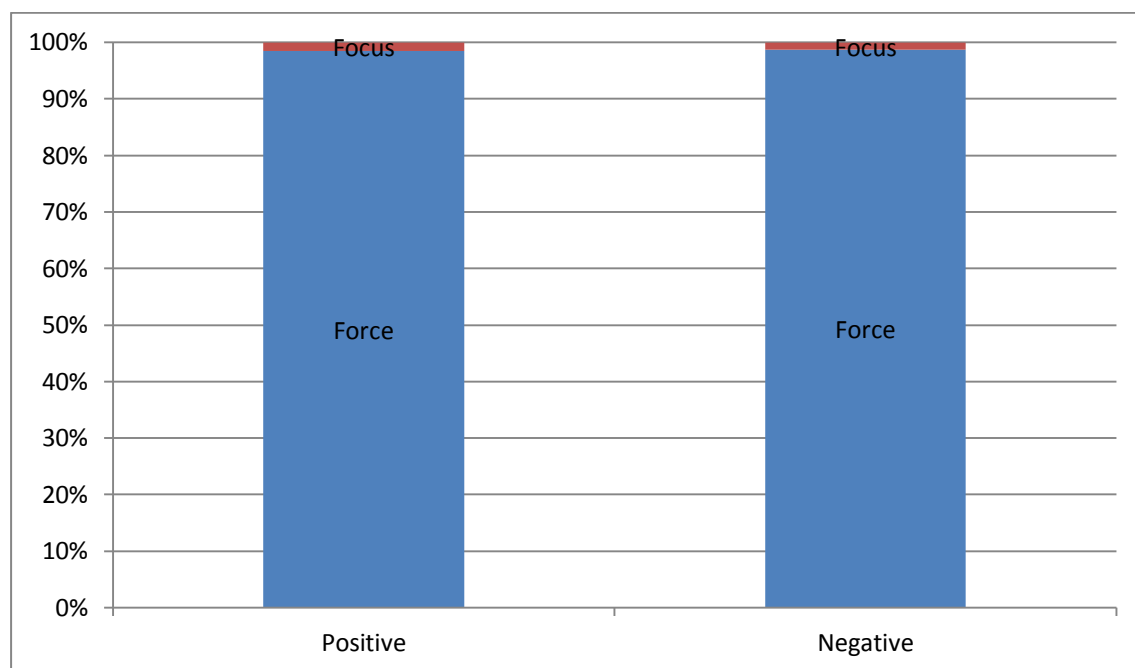


Figure 48. Distribution of Graduation in positive and negative reviews

The occurrences of Focus (1.46% vs. 1.31%) are so low that hardly any difference can be observed in any categorisation of Graduation. In this case, Force (98.54% vs. 98.69%) is again overwhelmingly exceeding Focus, which barely has any representation. Figure 50 below details the results for minor categories in Graduation.

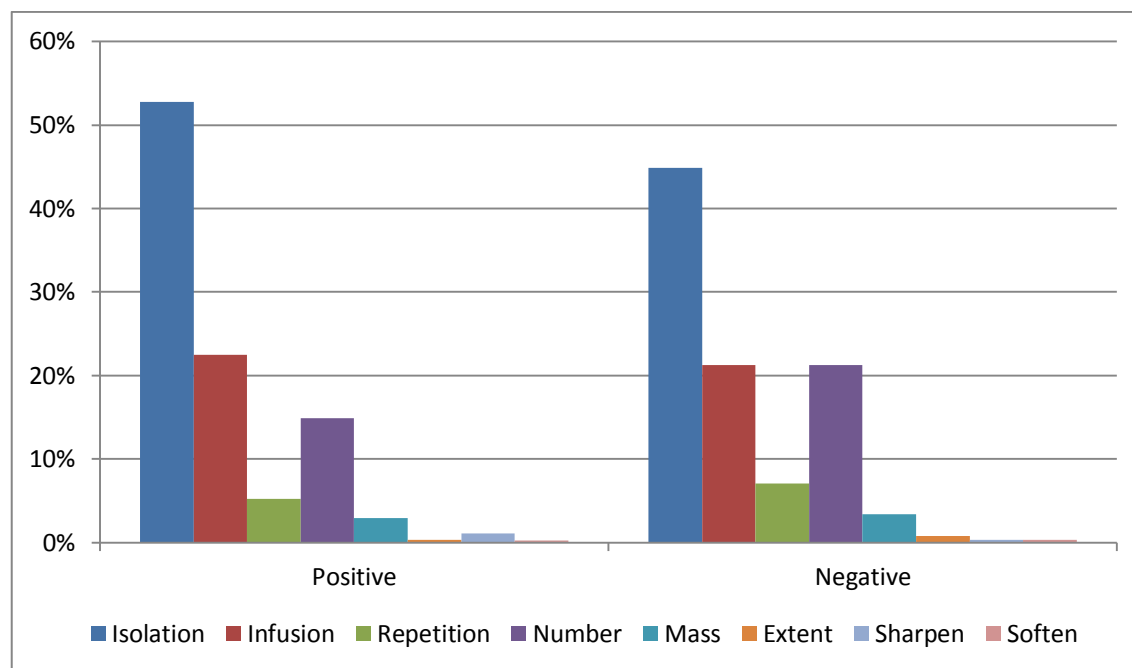


Figure 49. Distribution of Force and Focus in positive and negative reviews

Taking a deeper look at Graduation results, it can be seen that, although most of the categories do also share a similar representation, Isolation and Number do not, and show an inverse proportion: while Isolation is higher in positive comments than in negative texts (52.77% vs. 44.88%), Number has been more often annotated in the latter than in the former (14.87% vs. 21.26%). Authors rating items positively are more prone to emphasise the good qualities or feelings they have regarding that item by means of Isolation resources like *so*, *very*, etc., but annoyed critics are more likely to indicate number, typically a high one, such as *many* or *several*, to point out that a bad performance or a bad quality has been repeated, making things worse.

Nevertheless, the results obtained for the other three Force categories are virtually the same and do not show any tendency: Infusion (22.45% vs. 21.26%), Repetition (5.25% vs. 7.09%), Mass (2.92% vs. 3.41%), Extent (0.29% vs. 0.79%). Finally, the results for Focus are, as mentioned above, close to none: Sharpen (0.87% vs. 1.05%) and Soften (0.58% vs. 0.26%).

## 4.7. Summary and discussion

The results of the annotations in the larger set indicate interesting tendencies in the distribution of Appraisal tags in the bilingual reviews, although no statistically significant results (that is, with a  $P \leq 0.10$  value using a Chi-square test) were obtained. First, mobile applications reviews have been shown to be especially rich in Attitude tags, followed by Engagement, while Graduation tags occur much less frequently. This distribution reflects the communicative purpose of these texts, which is to present users' opinions on a given product. Therefore, the majority of the Appraisal tags are expressions of Attitude which assign a value to the item reviewed, or express someone's feelings related to that item. The need to engage other users in the comments is also reflected in the quite abundant use of Engagement tags in both the English and the Spanish reviews. Graduation tags, used to intensify or soften ideas, appear much less frequently in these reviews, indicating that users prefer other Appraisal strategies to convey their opinions on a given product.

As to the preferred tags from the extended tagsets, the reviews are rich in Appreciation expressions because they focus on the product, including its performance, qualities, effects, etc. while expressions of Affect and Judgement are less frequently comparatively. The most frequent subtypes of Appreciation tags are Valuation and Reaction, while Composition (how the object is composed) is less frequently used.

Affect tags are also common, but not as much as Appreciation. Affect deals with feelings and emotions. These meanings introduce the way the author or someone else feels in relation to the product reviewed and are the second-most common subtype of Attitude markers in the bilingual corpus. Their role in the reviews is usually supportive with respect to the role of Appreciation tags: if the qualities of the object itself are not enough to show why someone's opinion is the way it is, the expression of the users' feelings supports the emotional aspects of their opinion. The most common subtypes of Affect used in the larger corpus are (Un)Happiness and (Dis)Satisfaction. These include messages about how much users like (or dislike) the product or how satisfied and interested they are. Usually, authors tend to include Happiness expressions more often than Unhappiness elements, although Satisfaction and Dissatisfaction do not show such a clear distinction. (In)Security messages are not recurrent in these texts, so meanings related to fear, surprise, trust and the like are not frequently assigned to these products.



Judgement is the least used category in Attitude because it includes meanings used to evaluate behaviours. This kind of meanings is not commonly related to objects such as the products analysed here, or any objects at all. Despite this fact, and although it does not exceed Affect and Appreciation, more occurrences have been found than expected. More specifically, users were observed to focus on meanings related to Capacity and Propriety. Users classify a game's bugs as a theft, a character's behaviour as reprehensible or a singer as (un)talented, etc., thus increasing Judgement occurrences.

Engagement, as mentioned above, had the second highest rate after Attitude. Engagement is divided into two main categories: Expansion, which presents the author's voice as one in a range of possible voices, and Contraction, which delimits and denies other possible voices.

Expansion had the lowest results of the two, although some important categories are included under this label. In the first place, Expansion includes Entertain and Attribute. The latter had insignificant results in the corpus since authors share brief comments in which they go straight to the point and focus on their own opinion instead of including some else's, regardless of whether they agree or disagree with it. Entertain, however, showed more relevant results. The four Entertain types were used in the corpus: Epistemic resources are employed to modulate the author's certainty on their statements; Evidential markers indicate the empirical source of information (usually the sight, since the most frequent resource is *seem* and the like); Pseudo-Questions are commonly used linked to negative feelings of rejection and undesired surprise; and Deontic expressions indicate actions that either users have to do to get something in relation to the product or they give indications to readers or application designers.

Contraction types are highly frequent in the text, mainly due to Disclaim elements (Counter and Deny), which include common linguistic items such as conjunctions and negative particles. On the other hand, Proclaim resources (Concur, Pronounce and Endorse) are close to no occurrences at all. In the same way as Attribute spans were hardly covered in the corpus because authors focus on their own opinions instead of others', they are not interested in using Proclaim spans to provide a higher credibility or strength to their statements by means of someone else's authority or even reinforce their own authorial voice.

Finally, Graduation was the least used category, and its results are completely due to Force tags, since Focus ones are extremely scarce. These reviews do not take as their

common resource modulating the level of prototypicality of the nouns they use to name entities, but they do, and do a lot, intensify adjectives, verbs and indicate quantities for nouns. Thus, a product is not good but *very* good, a bug did not just happen, but happened *many* times, and they do not just like it, but like it *a lot*.

With respect to the comparison between English and Spanish texts, the latter use Attitude resources in a higher degree than the former, while this compensates its texts with higher Engagement and Graduation elements. Thus, Spanish texts draw on feelings and qualities much more frequently than English ones. The English reviews modulate their voice inside the text through Engagement as well as the words used by means of Graduation. Satisfaction meanings were preferred by the Spanish authors in a higher degree than the English ones, which focused on Happiness feelings. This may be a very subtle difference, but English writers have shown a strong attachment to the use of words like *love* and *like* for any kind of product while Spanish writers use *agradecer* [thank] or *esperar* [hope]. Similarly, Spanish writers have a higher interest in describing Capacity and Valuation than can be seen in the English texts, which lean more strongly on Normality, Veracity and Reaction meanings.

Regarding Engagement, the most noticeable aspects were the use of Epistemic and Pseudo-Questions in the case of English, and Deontic and Evidential markers in the case of Spanish. English reviews modulate certainty more extensively and are also more sarcastic by using rhetoric questions in their writings while Spanish ones are much more direct using obligation resources and basing their opinion on empirical sources. Following the same line, Counter elements are more profuse and varied in English, showing opposition and contrast, while Spanish writers are more direct by simply rejecting any other possibilities by means of Deny resources.

Conclusions on Graduation include a higher use of Repetition in the Spanish texts, mostly realised by the repetition of exclamation marks. No other differences have been found regarding the use of Focus (either Sharpen or Soften), due to the limited representation in both languages, nor in Intensification or Quantification.

Finally, it must be pointed out that the linguistic realisations used in both languages, presented in the most common spans in each of the three major axes, are very similar in English and in Spanish. Most of the linguistic items are actually shared as the most representative of each axis. Additionally, most of the most common linguistic items (typically, single words) were direct translations from one language into the other. They

were broad generic wild terms such as *good* or *bad*, *better*, *awesome*, etc. Therefore, authors mostly use a very simplistic language with highly recurrent words in common informal English and Spanish, instead of a field-specific terminology or jargon.

With respect to differences among products, some tendencies were observed that allowed the grouping of some products, indicating that each product has its own preferences regarding the use of Appraisal items. These can be summarised as follows: applications and games have a higher representation of positive Affect categories, such as Happiness, Security and Satisfaction, while books and films prefer the negative ones, namely Unhappiness and Insecurity. Music does also share the positive preference but with a visibly higher number of realisations of Happiness than the other four products. Capacity results are also comparable in applications, games and music, but no other similarities are found in the Attitude axis. Similarly, applications and games make a wider use of some Engagement types, namely Deontic elements, in contrast with the Epistemic resources that are so common in books and films. Regarding Graduation, again applications and games share their preference for Infusion spans while books and films pay more attention to Isolation markers. Finally, music joins applications and games when using Number resources.

These results establish two groups, one formed by applications and games, and a second one formed by books and films. Music shares some characteristics with both groups but does also show its own proper qualities, as the other four types do. These subtle similarities may point to a distinction between traditionally reviewed items, such as books and films, and the new ones, applications and games, since they may also share some inherent qualities: books and films have a story, characters, a director or writer, etc. while applications and games have performance issues. Music does not attach to any other the two types completely, and stands as a different group on its own.

With respect to differences between positive and negative reviews, the Attitude axis shows a preference for Judgement meanings in negative reviews that is not observed in positive texts. As mentioned above, Judgement meanings typically address morally incorrect behaviours (Propriety), since positive moral actions are taken for granted. Everybody must behave as rules say, so the positive accomplishment is not complimented, but negative behaviours are sanctioned and criticised in the texts. Positive Judgement realisations usually address Capacity meanings, such as talent, an adequate operation or improvements made in a product. Most common Affect meanings are

Happiness and Satisfaction in positive reviews and, unsurprisingly, Dissatisfaction and Unhappiness in the negative ones.

Engagement results show a higher use of Epistemic and Deontic resources in positive comments. For example, Authors introduce their opinions by means of spans like *I think* and do also recommend the product to other users through obligation meanings like *have to*. Negative texts use Pseudo-Questions and Evidential markers since they distance from the item by means of sarcastic questions or use verbs like *seem* to introduce a negative quality instead of stating it directly. Counter realisations were far more common than the other Disclaim type, Deny, in positive reviews, but they both obtained similar percentages in negative reviews. This is due to a higher use of negative elements in negative reviews, as can be expected, instead of a much lower use of Counter items.

Finally, Graduation differences include a higher interest in Isolation modifiers in positive reviews, while a more profuse use of Number items in negative ones. This means that words like *so* or *very* are typically attached to positive expressions like *good* instead of *bad*, while *many*, *some*, etc. are used when criticising a product.

## Chapter 5

# CONCLUSION AND FINAL REMARKS

The present study has aimed to answer four research questions presented in the introductory chapter. These research questions are copied below for convenience:

- (1) Is it possible to test the validity of aspects of Appraisal Theory empirically through corpus annotation? If so, are there features which need redefinition or reformulation to ensure high-quality bilingual (English-Spanish) annotation?
- (2) Are there language-independent Appraisal features which characterise mobile application reviews in English and Spanish? If so, which are these features?
- (3) Are there language-specific Appraisal features of mobile application reviews written in Spanish or in English?
- (4) Are there significant differences in the use of Appraisal features depending on the type of item reviewed and depending on the polarity features of the review?

The first research question addressed the nature of Appraisal Theory regarding its classification of language into categories. Appraisal Theory classifies evaluative language into three axis, namely Attitude (feelings and emotions), Engagement (involvement of not only the speaker's voice but also others' in the text) and Graduation (intensification or softening of both Attitude and Engagement expressions). This theory was applied to a

specific genre of evaluative texts: reviews published on the platform Play Store on products such as applications, games, books, films and music.

In order to validate empirically these categories as well as more delicate subcategories, three annotation experiments were designed. These experiments, also called agreement studies, measure the degree of agreement between independent coders, who are given specific guidelines and instructions for the task. The statistical measure used for this task was the *Kappa coefficient* (K), which provides a value for the degree of agreement between coders considering the actual agreement and the possible agreement by chance.

The first annotation experiment addressed the identification of markables in the reviews. Coders were asked to indicate the number of spans that were supposed to involve any Appraisal meaning. The results showed a very high agreement in the length of the spans selected, since only one span was annotated differently, but some disagreement occurred in the number of spans marked. The main reason for disagreement was the caution of one of the annotators, who usually overlooked some spans that were finally included as markables. These spans were long, typically phrases and sentences which are marginal examples of Appraisal, since most of the occurrences are individual terms.

Once the markables were identified and agreed upon between coders, a second annotation experiment was designed. This second experiment required the coders to label each of the spans previously marked with one tag from the core tagset. The core tagset consisted of three main categories extracted from the Appraisal Theory, i.e., Attitude, Engagement and Graduation, and major subtypes (Affect, Judgement, Appreciation, Expansion, Contraction, Force and Focus). A higher degree of agreement was obtained at this point than in the first annotation experiment. Most disagreement was found in Graduation, followed by Engagement and, finally, Attitude. Graduation types are used to modify both Attitude and Engagement, and this may lead to some confusion regarding the actual Graduation meaning of intensification or softening and the Attitude and Engagement meanings to which those changes are directed.

The third annotation experiment addressed specific tags from the extended tagset, including more delicate meanings from each of the main coarser tags. These included tags for meanings such as Happiness, Unhappiness, Security, Insecurity, Satisfaction, Dissatisfaction, Normality, Capacity, Tenacity, Veracity, Propriety, Reaction,

Composition, Valuation, Epistemic, Evidential, Pseudo-Question, Deontic, Acknowledge, Distance, Deny, Counter, Concur, Pronounce, Endorse, Sharpen, Soften, Isolation, Infusion, Repetition, Number, Mass and Extent. Moderate agreement was obtained when annotating specific subtypes of Attitude categories, where the confusion between Valuation and Reaction was especially noticeable. These categories involve the value and worthiness of the object reviewed, or a quality involving an impact on the user, respectively. These differences may be too subtle to be easily distinguished in any related terminology. The results for Engagement and Graduation showed an almost perfect agreement, and once these tags were selected, more delicate tags could also be distinguished by coders.

In conclusion, it can be stated that spans beyond the level of the word and, especially, the phrase, deserve further attention and study. Complex sentences may also involve Appraisal meanings, although these depend to a higher extent on human's interpretation. In addition, the coarser and more general types of Appraisal (Attitude, Engagement and Graduation) were validated, since a high level of agreement was reached. Interestingly, while Attitude was clearly identified and distinguished from other general tags, such as Engagement or Graduation, more delicate categories within Attitude were the most difficult to distinguish by coders. More specifically, the  $k$ -value obtained only reached a moderate agreement (0.49), mainly due to disagreement between Reaction and Valuation. The opposite was observed in Engagement and Graduation, whose axes were more confusing (they were mismatched twice and three times, respectively, in opposition to Attitude, which was mismatched only once), but inner categories obtained an almost perfect agreement. Regarding more delicate categories in particular, some specific types had a low or no representation in the corpus, such as Security, Insecurity, Tenacity, Concur, Endorse, Mass, Sharpen and Soften. Therefore, this type of texts may not be the most adequate genre to validate those specific categories, and they should be put to test again in a different genre with more occurrences.

The second research question focused on the language-independent Appraisal features which characterise mobile application reviews in English and Spanish. To investigate this question, a larger corpus of mobile application reviews was single-annotated by the author of this dissertation using the annotation scheme and the guidelines that were validated in the annotation experiments.

The results showed that these reviews are rich in Attitude expressions, especially Appreciation elements, because they focus on the product, including its performance, qualities, effects, etc. For this reason, the minor categories of Valuation (the quality itself) and Reaction (the effect the object produces on the user) were shown to have the highest number of occurrences, in comparison to Composition (how the object is composed) with very low number of occurrences. Affect is the second most common Attitude category because, after presenting all the items advantages and drawbacks, authors appeal to feelings. Judgement was the least used category because it includes meanings used to evaluate behaviours, and this kind of meanings is not commonly related to objects such as the products analysed here, or any objects at all. Regarding Engagement, authors have been shown to go straight to the point and focus on their own opinion instead of including some else's one, regardless they agree or disagree with it, so Attribute and Proclaim categories have a very low representation in the bilingual corpus. Finally, Graduation was mostly based on Force items. Thus, these reviews do not base their Graduation resources on modulating the level of prototypicality of the nouns they use to name entities, but on intensifying adjectives, verbs and indicating quantities for nouns.

The third research question addressed the contrastive analysis of Appraisal features between English and Spanish mobile application reviews in application stores. In order to answer this question, the results of bilingual annotation were divided into two language-specific datasets. Regarding the Spanish reviews, it was found that Attitude resources were highly preferred, especially Satisfaction, Capacity and Valuation, that is, feelings and qualities. As for Engagement tags, a high proportion of Deontic and Evidential markers was found, which indicates that the Spanish reviews are much more direct using obligation resources and basing their opinion on empirical sources. This straightforwardness is also reflected in the high frequency of the category of Deny, which simply rejects any other possibilities. Repetition realisations in Graduation are based on exclamation marks, which also manifest the intensity and directness of the Spanish texts.

The English reviews, on the other hand, show a higher proportion of Engagement and Graduation elements, modulating the author's voice inside the text. Interesting results were found for Epistemic and Pseudo-Questions, which are used to modulate certainty and show sarcasm respectively. This softening or modulation of the text was also realised by a profuse use of Counter elements, which show opposition and contrast. In the Attitude



axis, different categories from the Spanish ones stood out, namely Happiness, Normality, Veracity and Reaction.

Finally, the linguistic realisations for Attitude, Engagement and Graduation used in both languages showed a high degree of comparability. Actually, most of the most common linguistic items (typically, single words) were direct translations from one language into the other. Moreover, they were broad generic wild terms such as *good* or *bad*, *better*, *awesome*, etc. Therefore, authors mostly use a very simplistic language with highly recurrent words in common informal English and Spanish, instead of a field-specific terminology or jargon.

With respect to the fourth research question, the results of the bilingual annotation showed some differences depending both on the type of item reviews (applications, games, books, films and music) and polarity of the reviews (either positive or negative). The analysis of the reviews arranged by product did not reveal marked tendencies that could demonstrate the existence of clearly distinguished groups of reviews with different qualities, but some slight tendencies were observed.

On the one hand, applications and games had similar results in some specific subtypes of Attitude categories, such as Happiness, Security and Satisfaction (Affect), and Capacity (Judgement). Regarding Engagement, applications and games showed a wider use of this category, including a similar preference for Deontic items. In the Graduation axis, Infusion spans were noticeably important for these two products as well. On the other hand, books and films' results for negative elements such as Unhappiness and Insecurity were significant inside Attitude categories. The preference for Epistemic markers in Engagement did also show a coincidence between books and films, and the same happened with Isolation markers in Graduation. These subtle similarities may point to a distinction between traditionally reviewed items, such as books and films, and the new ones, applications and games, since they may also share some inherent qualities: books and films have a story, characters, a director or writer, etc. while applications and games have performance issues. Therefore, these preferences for some Appraisal categories may well be related to specific aspects of the item reviewed: for example, performance (Capacity) and actions required (Deontic) by the application or game, in opposition to a higher degree of involvement of the speaker in books and films (Engagement), which usually deal with more personal taste preferences.

Finally, music stands as a separated group, although some characteristics are shared more frequently with applications and games than with books and films. These include the positive preference of Attitude spans but with a significantly higher number of realisations of Happiness than any other products and comparable results in Capacity, as well as some coincidences with their results of Number in Graduation.

Another comparison was established between positive texts, those with a rating of 4 or 5 stars, and negative texts, rated with 1 or 2 stars. Negative texts are more prone to include Attitude negative feelings such as Unhappiness and Dissatisfaction, as well as Propriety evaluations criticising morally reprehensible behaviours. Regarding Engagement, they also use frequently Pseudo-Questions to show surprise or rejection of some fact and their opinions as grounded by Evidential markers. As can be expected, negative texts did also show a higher representation of Deny elements but they did also include Number items to indicate quantities in the Graduation axis. On the other hand, positive texts are characterised by Capacity appreciations as well as positive feelings such as Happiness and Satisfaction in the Attitude type. As far as Engagement markers are concerned, the annotation revealed a higher use of Epistemic and Deontic resources, either introducing their statements by warning the reader that it is just their subjective opinion or recommending the product through obligation meanings. Counter and Deny categories in Engagement were much more similar in positive texts, due to the lack of such a high proportion of Deny elements found in their counterpart. Finally, Graduation focused on the use of Isolation resources, thus meaning that emphasis is more typically linked to positive qualities. These differences point to usual aspects that are typically criticised or admired. For example, a good moral aspect will not be emphasised positively but it will be remarked if it does not adjust to the author's expectations. In the same way, positive performance (Capacity) will be praised if it satisfies the user but it is not so likely to be the focus of a bad review.

To conclude, with the work reported in this dissertation, I hope to have shed light into a number of interesting issues related to Appraisal Theory and its use in the characterisation of mobile application reviews. First, the creation of a large bilingual annotated dataset containing Appraisal features will hopefully contribute to the advancement of knowledge in areas related to Sentiment Analysis in NLP and the study of evaluation and interpersonal features of English and Spanish online reviews in the Corpus Linguistics

community. It is also expected that the empirical validation of the annotation scheme proposed in this dissertation contributes to the refinement and reformulation of Appraisal Theory within the SFL community and encourages further applied work to other genres and other languages. Second, the bilingual annotation process of Spanish and English mobile application reviews has provided an in-depth characterisation of this genre along Appraisal features, not only from the contrastive (English-Spanish) perspective, but also from the point of view of the type of item reviewed and the polarity assigned by the audience.

Future work will be focused on extending the current work to more specific lines of research in this area. Some of these lines include specific research on the realisation of Appraisal in long phrases and sentences, in order to find common validated features beyond readers' interpretations. Another interesting line of future research is the extension of the empirical validation of more delicate categories within Appraisal for which insufficient evidence was found in the current corpus. It would also be relevant to extend the current range of items reviewed to a wider range of products in order to find possible groups that share Appraisal features, thus confirming or diverging from the tendencies pointed out in this dissertation.



## REFERENCES

- Abdulla, N., N. Ahmed, M. Shehab, M. Al-Ayyoub, M. Al-Kabi and S. Al-Rifai. (2014). Towards Improving the Lexicon-Based Approach for Arabic Sentiment Analysis. *International Journal of Information Technology and Web Engineering*, 9(3). 55-71.
- Aone, C., M. Ramos-Santacruz and W. J. Niehaus. (2000). Assentor: An NLP-based solution to e-mail monitoring. *Proceedings of the 12th Innovative Applications of Artificial Intelligence Conference (IAAI-2000)*, Austin, TX, August 1–3. 945–950.
- Bakhtin, M. (1986). *Speech genres and other late essays*, Vern W. McGee (trans.). Austin: University of Texas Press.
- Balahur, A., Z. Kozareva and A. Montoyo. (2009). Determining the polarity and source of opinions expressed in political debates. *Proceedings of the Tenth International Conference on Intelligent Text Processing and Computational Linguistics*. 468-480.
- Banfield, A. (1982). *Unspeakable Sentences*. Boston: Routledge and Kegan Paul.
- Becker, A. (2009). Modality and Engagement in British and German political interviews. *Languages in Contrast*, 9(1). 5-22.
- Bhatia, V. K. (2004). *Worlds of Written Discourse: A Genre-Based View*. London: Continuum International.
- Bloom, K., N. Garg and S. Argamon. (2007). Extracting Appraisal Expressions. *Proceedings of NAACL HLT*. Rochester. 308–315
- Bloom, K., S. Stein and S. Argamon. (2007). Appraisal Extraction for News Opinion Analysis at NTCIR-6. *Proceedings of NTCIR-6 Workshop Meeting*. Tokyo, Japan.

- Boldrini, E., J. Fernández, J.M. Gómez and P. Martínez-Barco. (2009). Machine learning techniques for automatic opinion detection in non-traditional textual genres. *1<sup>st</sup> workshop on Opinion Mining and Sentiment Analysis (WOMSA)*.
- Callaham, J. (2015). Google says there are now 1.4 billion active Android devices worldwide. *Android Central*. (<http://www.androidcentral.com/google-says-there-are-now-14-billion-active-android-devices-worldwide>)
- Carbonell, J. G. (1979). *Subjective Understanding: Computer Models of Belief Systems*. Ph.D. thesis and Technical Report no. 150, Department of Computer Science, Yale University, New Haven, CT.
- Carretero, M. and M. Taboada. (2009). Contrastive analyses of Evaluation in text: Key issues raised by the application of Appraisal Theory to corpora of consumer-generated product reviews in English and Spanish. *I Congreso Internacional de Lingüística de Corpus*. Murcia, Spain.
- Carretero, M. and M. Taboada. (2010a). Products, consumers and evaluation: a proposal of solutions to problematic issues of Attitude in English and Spanish consumer reviews. *22nd European Systemic Functional Linguistics Conference and Workshop*. University of Primorska, Slovenia.
- Carretero, M. and M. Taboada. (2010b). The annotation of Appraisal: How Attitude and epistemic modality overlap. *4th International Conference on Modality in English (ModE4)*. Madrid, Spain.
- Carretero, M. and M. Taboada. (2011). Annotating Appraisal: contrastive issues raised in the analysis of consumer reviews in English and Spanish. *38<sup>th</sup> International Systemic Functional Linguistic Congress*. Lisbon, Portugal.
- Carretero, M. and M. Taboada. (2014). Graduation within the scope of Attitude in English and Spanish consumer reviews of books and movies. In G. Thompson and L. Alba-Juez (eds.), *Evaluation in Context*. Amsterdam: John Benjamins. 221-239.
- Carretero, M. and M. Taboada. (2015). The annotation of Appraisal: How Attitude and Epistemic Modality overlap in English and Spanish consumer reviews. In J.R. Zamorano (ed.), *Thinking Modally: English and Contrastive Studies on Modality*. Newcastle: Cambridge Scholars. 429-270.
- Carrillo de Albornoz, J., L. Plaza, P. Gervás, and A. Díaz. (2011). A joint model of feature mining and sentiment analysis for product review rating. *Proceedings of ECIR*. 55–66.
- Chaovalit, P. and L. Zhou. (2005). Movie review mining: A comparison between supervised and unsupervised classification approaches. *Proceedings of the Hawaii International Conference on System Sciences (HICSS)*.
- Chevalier J. A. and D. Mayzlin. (2006). The effect of word of mouth on sales: Online book reviews. *Journal of Marketing Research*, 43 (August). 345–354.

- Coffin, C. (1997). Constructing and giving value to the past: An investigation into secondary school history. In F. Christie and J. R. Martin (Eds.), *Genre and Institutions: Social Processes in the Workplace and School*. London: Cassell. 196-230.
- Conrad, A. J. Wiebe, and R. Hwa. (2012). Recognizing arguing subjectivity and argument tags. Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics, ExProM '12. 80-88.
- Cruz, F.L, J.A. Troyano, F. Enríquez, F.J. Ortoga. (2008). Clasificación de documentos basada en la opinión: Experimentos con un corpus de críticas de cine en español. *Boletín de la Sociedad Española para el Procesamiento del Lenguaje Natural*, 41. 73-80.
- de Jong, I. K., and C. Burgers. (2013). Do consumer critics write differently from professional critics? A genre analysis of online film reviews. *Discourse, Context & Media*, 2(2). 75-83.
- Deng, L., Y. Choi and J. Wiebe. (2013). Benefactive/malefactive event and writer attitude annotation. *ACL* 2. 120-125.
- Deng, L. and J. Wiebe. (2015). MPQA 3.0: An Entity/Event-Level Sentiment Corpus. *Human Language Technologies: The 2015 Annual Conference of the North American Chapter of the ACL*. 1323-1328.
- Di Bari M., S. Sharoff and M. Thomas. (2015). A manually-annotated Italian corpus for fine-grained sentiment analysis. *Proceedings of the 2<sup>nd</sup> Italian Conference on Computational Linguistics (CliCit)*. 105-109.
- Dligach, D., R.D. Nielsen and M. Palmer. (2010). To annotate more accurately or to annotate more. *Proceedings of the Fourth Linguistic Annotation Workshop*. 64{72}, 55, 66, 132, 337, 380, 389.
- Dong, A. (2006). How am I doing? The language of Appraisal in design. In J.S. Gero (ed.), *Design Computing and Cognition*. Dordrecht: Springer. 385-404.
- Eggins, S. (1994/2004). *An Introduction to System Functional Linguistics*. London: Pinter Publishers.
- Guzmán, E. and W. Maalej. (2014). How do users like this feature? A fine grained sentiment analysis of app reviews. *Requirements Engineering Conference (RE)*.IEEE 22nd International.
- Hovy, E.H. and J. Lavid. (2010). Towards a ‘Science’ of Corpus Annotation: A new Methodological Challenge for Corpus Linguistics. *International Journal of Translation*, 22(1). 13-36.
- Hommerberg, C. and A. Don. (2015). Appraisal and the language of wine appreciation: A critical discussion of the potential of the Appraisal framework as a tool to analyse specialised genres. *Functions of Language*, 22(2). 161-191.

- Hood, S. and J. Martin. (2005). Invoking attitude: the play of graduation in appraising discourse. *Revista Signos*, 38 (58), 295-200.
- Hu, M. and B. Liu. (2004). Mining and summarizing customer reviews. *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*. 168-177.
- Hunston, S. (2003). Frame, phrase or function: a comparison of frame semantics and local grammars. In Archer et al (eds.), *Proceedings of the Corpus Linguistics 2003 Conference*. UCREL Technical Papers, vol. 16. Lancaster: UCREL.
- Hunston, S. (2010). *Corpus Approaches to Evaluation. Phraseology and Evaluative Language*. London: Routledge.
- Hunston, S. and G. Thompson (eds.). (2000). *Evaluation in text: authorial stance and the construction of discourse*. Oxford [etc.]: Oxford University Press.
- Islam, M.R. (2014). Numeric rating of Apps on Google Play Store by sentiment analysis on user reviews. *Electrical Engineering and Information & Communication Technology (ICEEICT). International Conference*. IEEE.
- Kennedy A. and D. Inkpen. (2006). Sentiment classification of movie reviews using contextual valence shifters. *Computational Intelligence*, 22. 110–125.
- Khalid, H. (2013). On identifying user complaints of iOS apps. *Software Engineering (ICSE), 35th International Conference*. IEEE.
- Khalid, M., M. Asif and U. Shehzaib. (2015). Towards improving the quality of mobile app reviews. *International Journal of Information Technology and Computer Science (IJITCS)*, 7(10). 35-41.
- Kimura, H. (2014). 7 types of Google Play App reviewers and how to respond to them. *Sensor Tower*. (<https://sensortower.com/blog/7-types-of-google-play-app-reviewers-and-how-to-respond-to-them>)
- Konerek, P. and M. Šimko. (2013). Sentiment analysis on microblog utilizing appraisal theory. *World Wide Web*. 1-21.
- Kontopoulos, E., C. Berberidis, T. Dergiades and N. Bassiliades. (2013). Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*, 40(10). 4064-4074.
- Lavid, J. (2012). Corpus analysis and annotation in CONTRANOT: Linguistic and Methodological Challenges. In I. Moskowich and B. Crespo (eds.), *Encoding the past, decoding the future: corpora in the 21<sup>st</sup> century*. Cambridge: Cambridge Scholars. 205-220.
- Lavid, J. (2017). Annotating complex linguistic features in bilingual corpora: The case of MULTINOT. In T. Declerck and S. Kübler (eds.), *Proceedings of the Workshop on Corpora in the Digital Humanities (CDH 2017)*. IN, USA: Bloomington. Available online at <http://ceur-ws.org/Vol-1786/>.



- Lavid, J., J. Arús, M. Carretero, L. Moratón and J.R. Zamorano. (2014). Contrastive corpus annotation in the CONTRANOT Project: Issues and problems. In M.A. Gómez González, F. J. Ruiz de Mendoza Ibáñez, F. González-García and A. Downing (eds.), *The Functional Perspective on Language and Discourse: Applications and implications*. Amsterdam: John Benjamins. 57-86.
- Lavid, J., J. Arús, B. DeClerk, V. Hoste. (2015). Creation of a High-quality, Register-diversified Parallel (English-Spanish) Corpus for Linguistic and Computational Investigations. *7<sup>th</sup> International Conference on Corpus Linguistics: Current Work in Corpus Linguistics: Working with Traditionally-conceived Corpora and Beyond (CILC)*. Procedia – Social and Behavioral Sciences, Volume 198. 249-256.
- Lavid, J., M. Carretero and J.R. Zamorano. (2016). Contrastive Annotation of Epistemicity in the MULTINOT Project: Preliminary Steps.
- Leech, G. (1997). Introducing Corpus Annotation. In R. Garside, G. Leech and A. McEnery (eds.), *Corpus Annotation: Linguistic Information from Computer Text Corpora*. London: Longman.
- Li, N. and D.D. Wu. (2009). Using text mining and sentiment analysis for online forums hotspot detection and forecast. *Decision Support Systems* 48 (2). 354-368.
- Liu, Bing. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool.
- Love, K. (2006). APPRAISAL in online discussions of literary texts. *Text and Talk*, 26 (2), 217-244.
- Macken-Horarik, M. (2003). APPRAISAL and the special instructiveness of narrative. *Text*, 23 (2), 285-312
- Macken-Horarik, M. and A. Isaac. (2014). Appraisal Appraisal. In G. Thompson and L. Alba-Juez (eds.), *Evaluation in Context*. Amsterdam/Philadelphia: John Benjamins.
- Marín-Arrese, J.I. and B. Núñez Perucha. (2006). Evaluation and Engagement in Journalistic Commentary and News Reportage. *Revista Alicantina de Estudios Ingleses*, 19. 225-248.
- marketsandmarkets.com. (2010). World Mobile Applications Market - Advanced Technologies, Global Forecast (<http://www.marketsandmarkets.com/Market-Reports/mobile-applications-228.html>).
- Martin, J.R. (1996). Evaluating disruption: symbolising theme in junior secondary narrative, in R. Hasan and G. Williams (eds.), *Literacy in Society*, 124-171. London: Longman.
- Martin, J.R. (2000). Beyond exchange: Appraisal systems in English. In S. Hunston and G. Thompson (Eds.), *Evaluation in Text: Authorial Distance and the Construction of Discourse*. 142-175. Oxford: Oxford University Press.

- Martin, J.R. (2001). Writing history: construing time and value in discourses of the past. In C. Colombi and M. Schleppergrell (eds.), *Developing Advanced Literacy in First and Second Languages*. Mahwah, N.J.: Erlbaum.
- Martin, J.R. & P.R.R. White (2005). *The Language of Evaluation. Appraisal in English*. New York: Palgrave.
- McCabe, A. (2007). Textual and interpersonal differences between a news report and an editorial. *Revista Alicantina de Estudios Ingleses*, 20. 139-156.
- Miller, C. (1984/1994). Genre as social action. In A. Freedman and P. Medway (Eds.), *Genre and the new rhetoric*. 23-42. London: Taylor and Francis.
- Molina-González, M.D., E. Martínez-Cámara, M.T. Martín-Valdivia and J.M. Perea-Ortega. (2013). Semantic orientation for polarity classification in Spanish reviews. *Expert Systems with Applications*, 40. 7250-57.
- Mora, N. (2011). *Annotating Expressions of Engagement in Online Book Reviews: A Contrastive (English-Spanish) Corpus Study for Computational Processing*. <http://eprints.ucm.es/13754/>
- Moreno-Ortiz, A. and J. Fernández Cruz. (2015). Identifying polarity in financial texts for Sentiment Analysis: A Corpus-based Approach. *Procedia: Social and Behavioral Sciences*. Vol. 198. 330-338.
- Moreno-Ortiz, A. and C. Gómez Pascual. (2016). Extracting domain-Specific features for Sentiment Analysis using simple NLP techniques: Running Shoes reviews. *EPiC Series in Language and Linguistics*. 298-307.
- Moreno-Ortiz, A. and C. Pérez Hernández. (2013). *Lexicon-based Sentiment Analysis of Twitter messages in Spanish*. *Procesamiento del Lenguaje natural*, 50. 93-100.
- Moreno-Ortiz, A., C. Pérez Hernández and R. Hidalgo García. (2011a). Domain-neutral-Linguistically-motivated Sentiment Analysis: a performance evaluation. *Actas del XXVII Congreso de la Sociedad Española para el Procesamiento del Lenguaje Natural*. 361-369.
- Moreno-Ortiz, A., C. Pérez Hernández and R. Hidalgo García. (2011b). Utilización de corpora textuales para la extracción de modificadores contextuales de Valencia para tareas de Análisis de Sentimiento. *Actas del 3<sup>er</sup> Congreso Internacional de Lingüística de Corpus. Tecnologías de la Información y las Comunicaciones: Presente y Futuro en el Análisis de Corpus*. 847-856.
- Moreno-Ortiz, A., F. Pineda Castillo and R. Hidalgo García. (2010). Análisis de valoraciones de usuario de hoteles con Sentitext: un sistema de análisis de sentimiento independiente del dominio. *Procesamiento de Lenguaje Natural*, 45. 31-39.
- O'Mahony, M. P. and B. Smyth. (2009). Learning to recommend helpful hotel reviews. *Proceedings of the third ACM conference on Recommender systems*.

- Ortigosa, A., J.M. Martín and R.M. Carro. (2014). Sentiment analysis in Facebook and its application to e-learning. *Computers in Human Behavior* 31. 527-541.
- Page, R.E. (2003). An analysis of APPRAISAL in childbirth narratives with special consideration of gender and storytelling style. *Text*, 23 (2), 211-237
- Painter, C. (2003). Developing attitude: an ontogenic perspective on APPRAISAL. *Text* 23(2). 183-210.
- Pang, B., L. Lee and S. Vaithyanathan. (2002). Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of the 2002 Conference on Empirical Methods in Natural Language Processing*, Philadelphia, PA, EEUU.
- Pang, B. and L. Lee. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2). 1-135.
- Poria, S., E. Cambria, G. Winterstein and G.-B. Huang. (2014). Sentiment patterns: Dependency-based rules for concept-level sentiment analysis. *Knowledge-Based Systems*, 69. 45-63.
- Pounds, G. (2010). Attitude and subjectivity in Italian and British hard-news reporting: The construction of a culture-specific 'reporter' voice. *Discourse Studies*, 12(1). 106-137.
- Prabowo R. and M. Thelwall. (2009). Sentiment analysis: A combined approach. *Journal of Informetrics* 3 (2), 143-157.
- Quirk, R., S. Greenbaum, G. Leech and J. Svartvik. (1985). *A Comprehensive Grammar of the English Language*. New York: Longman.
- Sack, W. (1995). Representing and recognizing point of view. *Proceedings of the AAAI Fall Symposium on AI Applications in Knowledge Navigation and Retrieval*, Cambridge, MA. 152.
- Schumaker, R.P., Y. Zhang, C.-N. Huang and H. Chen. (2012). Evaluating sentiment in financial news articles. *Decision Support Systems* 53 (3). 458-464.
- Socher, R. A. Perelygin, J.Y. Wu, J. Chuang, C.D. Manning, A.Y. Ng and A. Potts. (2013). Recursive deep models for semantic compositionality over a sentiment treebank. *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*. 1631-1642.
- Somasundaran, S., J. Ruppenhofer and J. Wiebe. (2008). Discourse Level Opinion Relations: An Annotation Study. *Proceedings of the 9th SIGdial Workshop on Discourse and Dialogue*. Columbus. 129-137.
- Sano, M. (2011). Reconstructing the English system of Attitude for the application to Japanese: an exploration for the construction of a Japanese Dictionary of Appraisal. Paper presented at the 38<sup>th</sup> International Systemic Functional Congress.

- Sorensen A.T. and S.J. Rasmussen. (2004). Is any publicity good publicity? A note on the impact of book review. *NBER Working paper*, Stanford University.
- Swales, J. M. (1981). *Aspect of article introductions* (Aston ESP Research Report 1). Birmingham, England: University of Aston in Birmingham, Language Studies Unit.
- Swales, J. M. (1986). A genre-based approach to language across the curriculum. In M. L. Tickoo (Ed.), *Language across the curriculum*. Singapore: Regional English Language Center. 10-22.
- Swales, J.M. (1990). *Genre analysis: English in academic and research settings*. Cambridge: Cambridge University Press.
- Taboada, M. (2008). The SFU Review Corpus. Simon Fraser University.
- Taboada, M. (2009). Comparable corpora for cross-linguistic sentiment analysis. *Conference of the American Association for Corpus Linguistics*. Edmonton, Alberta.
- Taboada, M. (2011). Stages in an online review genre. *Text and Talk*, 31(2). 247-269.
- Taboada, M. (2016). Sentiment analysis: An overview from linguistics. *Annual Review of Linguistics*, 2. 325-347.
- Taboada, M. and J. Grieve. (2004). Analyzing appraisal automatically. *Proceedings of AAAI Spring Symposium on Exploring Attitude and Affect in Text (AAAI Technical Report SS-04-07)*. Stanford University, CA, EEUU. 158-16.
- Taboada, M., J. Brooke and M. Stede. (2009). Genre-Based Paragraph Classification for Sentiment Analysis. *Proceedings of 10<sup>th</sup> Annual SIGDIAL Conference on Discourse and Dialogue*. 62-70.
- Taboada, M., J. Brooke, M. Tofiloski, K. Voll and M. Stede. (2011). Lexicon-Based Methods for Sentiment Analysis. *Computational Linguistics*, 37(2). 267-307.
- Taboada, M. M. Carretero and J. Hinnel. (2014). Loving and hating the movies in English, German and Spanish. *Languages in Contrast*, 14(1). 127-161.
- Tan, S. and J. Zhang. (2008). An empirical study of sentiment analysis for Chinese documents. *Expert Systems with Applications*, 34(4). 2622-2629.
- Tang, D., B. Qin and T. Liu. (2015). Deep learning for sentiment analysis: successful approaches and future challenges. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 5(6). 292-303.
- Terveen, L., W. Hill, B. Amento, D. McDonald and J. Creter. (1997). Building task-specific interfaces to high volume conversational data. *Proceedings of the Conference on Human Factors in Computing Systems (CHI-97)*, Los Angeles, April 18-23. 226-233.
- Thompson, G. and L. Alba-Juez. (2014). *Evaluation in Context*. Amsterdam: John Benjamins.

- Thomson, E.A., P.R. White and P. Ketley. (2008). "Objectivity" and "hard news" reporting across cultures: comparing the news report in English, French, Japanese and Indonesian journalism. *Journalism Studies*, 9(2). 212-228.
- Titov, I. and R.T. McDonald. (2008). A joint model of text and aspect ratings for sentiment summarization. *ACL*, 8. 308-316.
- Turney, P. (2002). Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL-2000)*, Philadelphia, July 7–12. 417–424.
- Vasa, R., et al. (2012). A preliminary analysis of mobile app user reviews. *Proceedings of the 24th Australian Computer-Human Interaction Conference*. ACM.
- Vian Jr., O. (2008). Appraisal System in Brazilian Portuguese: Resources for Graduation. *Odense Working Papers in Language and Communication*, 29. 808-824.
- Viera, A.J. and J.M. Garret. (2005). Understanding Interobserver Agreement: The Kappa Statistic. *Family Medicine*, 37(5), 360-3.
- Vilares, D., M.A. Alonso, C. Gómez-Rodríguez. (2015). A syntactic approach for opinión mining on Spanish reviews. *Natural Language Engineering*, 21(1). 139-163.
- Wang, J. and A. Dong. (2007). How am I doing 2: Computing the language of Appraisal in design. *International Conference on Engineering Design*, Paris.
- White, P.R.R. (1998). *Telling Media Tales: The News Story as Rhetoric*. Unpublished Ph.D. dissertation. Sydney: University of Sydney
- White, P.R.R. (2002) Appraisal. In J. Verschueren et al. (eds.). *Handbook of pragmatics*, 1–27. Amsterdam: Benjamins.
- White, P.R.R. (2003). Beyond modality and hedging: A dialogic view of the language of intersubjective stance. *Text*, 23 (2), 259–284.
- White, P.R.R. (2004). Subjectivity, Evaluation and Point of View in Media Discourse. In C. Coffin, A.Hewings and K. O'Halloran (eds.), *Applying English Grammar*. London: Hodder Arnold.
- White, P.R.R. (2011). Language as dialog – Engagement and the analysis of interpersonal positioning in student writing and media commentary (Workshop). *38th International Systemic Functional Congress (ISFC38): Negotiating difference: languages, metalanguages, modalities, cultures*. Lisbon, Portugal.
- Whitelaw, C. et al. (2005). Using appraisal taxonomies for sentiment analysis. In *Proceedings of MCLC-05, the 2nd Midwest Computational Linguistic Colloquium*, Colombia, EEUU.

- Whitelaw, C., N. Garg and S. Argamon. (2005). Using appraisal groups for sentiment analysis. *Proceedings of the 14<sup>th</sup> ACM international conference on Information and Knowledge management*. 625-631.
- Wiebe, J. (1994). Tracking point of view in narrative. *Computational Linguistics*, 20(2). 233–287.
- Wiebe, J., T. Wilson, R. Bruce, M. Bell, and M. Martin. (2004). Learning subjective language. *Computational linguistics*, 30(3). 277–308.
- Wiebe, J., T. Wilson and C. Cardie. (2005). Annotating Expressions of Opinions and Emotions in Language. *Language Resources and Evaluation*, 39(2-3), 165-210.
- Wilson, T. (2007). *Fine-grained Subjectivity and Sentiment Analysis: Recognizing the Intensity, Polarity, and Attitudes of private states*. Ph.D. thesis, Intelligent Systems Program, University of Pittsburgh.
- Wilson, T. and J. Wiebe. (2005). Annotating attributions and private states. *Proceedings of the Workshop on Frontiers in Corpus Annotation II: Pie in the Sky*. Ann Arbor. 53-60.
- Wilson, T., J. Wiebe and P. Hoffmann. (2009). Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis. *Computational Linguistics* 35 (3). 399-433.
- Ying, Y. (2012). An overview of the discourse functions of graduation within Appraisal. *Foreign Language Research*, 1, 84-88.
- Yu, H. and V. Hatzivassiloglou. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. *Proceedings of the 203 conference on Empirical methods in natural language processing*. 129-136.
- Zhuang, L., F. Jing, X. Y. Zhu, and L. Zhang. (2006). Movie review mining and summarization. *Proceedings of the ACM SIGIR Conference on Information and Knowledge Management (CIKM)*.

## APPENDICES





## Appendix A. CURRICULUM VITAE

### Education

---

9/2015 to present	<b>MA in <i>Computer Systems and Languages</i> (UNED)</b> <ul style="list-style-type: none"><li>• This MA focuses on the access, exploration and analysis of large quantities of textual information and the WWW.</li></ul>
9/2012 to 2017	<b>PhD in <i>English Linguistics</i> (Universidad Complutense de Madrid)</b> <ul style="list-style-type: none"><li>• Thesis: <i>Annotating Appraisal in English and Spanish: a Contrastive Study for Linguistic and Computational purposes.</i></li></ul>
10/2012 to 9/2013	<b>MA in <i>Teacher Training: Secondary and Upper Secondary Education, Vocational Training and Language Teaching</i> (Universidad Complutense de Madrid)</b> <ul style="list-style-type: none"><li>• Thesis: <i>The Use of Construction Grammar in ESL: A Case Study.</i></li></ul>
9/2010 to 6/2013	<b>BA in <i>Spanish Studies</i> (Universidad Complutense de Madrid)</b> <ul style="list-style-type: none"><li>• Major in Spanish Grammar.</li></ul>
10/2010 to 9/2011	<b>MA in <i>English Linguistics: New Applications and International Communication</i> (Universidad Complutense de Madrid)</b> <ul style="list-style-type: none"><li>• Thesis: <i>Annotating Expressions of Engagement in Online Book Reviews: A Contrastive (English-Spanish) Corpus Study for Computational Processing.</i></li></ul>
9/2005 to 6/2010	<b>BA in <i>English Studies</i> (Universidad Complutense de Madrid)</b> <ul style="list-style-type: none"><li>• Major in British Literature.</li><li>• Major in Applied Linguistics.</li></ul>

### Professional experience

---

9/2016 to present	<b>Co-Founder (Idiom Coach S.L.)</b> <ul style="list-style-type: none"><li>• Proofreading online services for ESP (English for Specific Purposes) and GE (General English).</li></ul>
----------------------	---

<b>4/2016 to present</b>	<b>Course syllabus and content editor (Lexgal Formación)</b> <ul style="list-style-type: none"> <li>· Design of ESP courses and GE.</li> </ul>
<b>9/2014 to present</b>	<b>Teacher of English as Foreign Language and Spanish Language and Literature (Escuela Politécnica Giner)</b> <ul style="list-style-type: none"> <li>· Teacher of English as foreign language and Spanish language and literature with students from 16 years old onward.</li> </ul>
<b>9/2013 to 8/2014</b>	<b>Software, usability and localization tester (Samsung Electronics Iberia)</b> <ul style="list-style-type: none"> <li>· Tester of the application S Voice, based on speech recognition and processing. Identification, classification and analysis of errors in S Voice linguistic processing and output, including phonetic, grammatical and semantic mistakes. Support on different tasks, such as translation and localization and devices' texts and user guides reviewing. Development of grammars with Eclipse in Augmented Backus-Naur Form (ABNF).</li> </ul>
<b>9/2009 to 6/2013</b>	<b>Proofreader and translator English-Spanish, Spanish-English (Jarpyo Editores SA)</b> <ul style="list-style-type: none"> <li>· Correction and translation of Spanish and English texts.</li> </ul>
<b>3/2013 to 5/2013</b>	<b>Corpus annotator (UNED)</b> <ul style="list-style-type: none"> <li>· Corpus annotator in the project "LiMoSINe (FP7-ICT-2011-7)–Linguistically Motivated Semantic Aggregation Engines". Semantic and polarity annotation and classification of opinionated texts with in-house tools.</li> </ul>
<b>12/2009 to 6/2011</b>	<b>English-Spanish translator (Universidad Complutense de Madrid)</b> <ul style="list-style-type: none"> <li>· English-Spanish translator in the project "SIIM (FFI2008-05388/FISO 2009-2011) – Studies on Intermediality and Intercultural Mediation".</li> </ul>
<b>12/2009 to 6/2011</b>	<b>Internship at the English Department (Universidad Complutense de Madrid)</b> <ul style="list-style-type: none"> <li>· Participation in the organization of conferences, workshops and other administrative tasks supporting professors and the head of the department.</li> </ul>

## Publications and presentations at conferences and courses

---

- Mora, N. (2014). Evidential and Epistemic expression in English and Spanish online reviews. Paper presented at EMEL (International Conference on Evidentiality and Modality in European Languages). Universidad Complutense de Madrid, Spain, 6-8 October 2014.
- Mora, N. (2013). Schopenhauer's determinism in Calderón's *La vida es sueño* and *El médico de su honra*. In A. Kuzmanovic Jovanovic, J. Stojanovic, J. Filipovic y J. Rajic (eds.), *Estudios Hispánicos en el siglo XXI*. University of Belgrade. <http://www.fil.bg.ac.rs/wp-content/uploads/obavestenja/iberijske/ehes21/02Mora%20Lopez.pdf>
- Mora, N. (2012). Modality in four texts of the independence of the English and Spanish colonies. *Diálogo de la Lengua* IV, 39-52. [http://www.dialogodelalengua.com/articulo/pdf/4/3\\_mora\\_dl\\_2012.pdf](http://www.dialogodelalengua.com/articulo/pdf/4/3_mora_dl_2012.pdf)
- Mora, N. (2012). The Underlying Feminist Message in Susan Glaspell's *Trifles*: A Hallidayan Approach. Theoretical and practical session. Course *Studies on Functional Linguistics Applied to English* (MA in English Linguistics: New Applications and International Communication). Universidad Complutense de Madrid.
- Mora, N. and Y. Rosolovska. (2011). The Underlying Feminist Message in *Trifles*: A Hallidayan Approach. Paper presented at 2<sup>nd</sup> Conference of Young Researches on Anglophone Studies. Universidad de Salamanca, Spain. 6-8 October 2011.
- Mora, N. (2011). *Annotating Expressions of Engagement in online Book reviews: A Contrastive (English Spanish) Corpus Study for Computational Processing*. MA Thesis (<http://eprints.ucm.es/13754/>)



## Appendix B. ENGLISH AND SPANISH SUMMARY

### English summary

Evaluation, opinion and subjectivity are related phenomena which are currently receiving attention both in the linguistic and the computational communities (Wiebe et al. 2004; Pang et al. 2002; Taboada 2009; Taboada and Grieve 2004; Whitelaw et al. 2005; Bloom, Garg and Argamon 2007; Bloom, Stein and Argamon 2007; Pang and Lee 2008, *inter alia*) under different theories and approaches.

One of the most influential theories that deals with the phenomenon of evaluation and subjectivity is the ‘Appraisal’ framework, a linguistic approach for ‘exploring, describing and explaining the way language is used to evaluate, to adopt stances, to construct textual personas and to manage interpersonal positionings and relationships’ (see Appraisal Homepage at <http://grammatics.com/appraisal/>). Appraisal Theory has been developed by Martin and White in Australia (see Martin 2000; Martin and White 2005; White 2003; White 2004, *inter alia*) and proposes that linguistic expressions of evaluative meanings such as emotion, attitude and opinion can be divided into three different axes: Attitude, Engagement and Graduation.

As a result of the great impact of this theory, considerable research has been devoted to applying and expanding the theory in different ways, focusing mainly on the area of Attitude (e.g. Macken-Horarik 2003; Painter 2003; Page 2003; White 2004), with less attention being paid to the areas of Engagement and Graduation, which require further study. Appraisal Theory has also been applied to other languages and to different genres (Sano 2011; Vian Jr. 2008; Pounds 2010; Becker 2009; Thomson, White and Kitley 2008; Taboada, Carretero and Hinnel 2014).

However, to date, there is no large-scale cross-linguistic work in the computational or in the linguistic communities which has set out the task to validate the different features of Appraisal Theory empirically through corpus annotation, and to investigate its application to a relatively new genre, namely, mobile application reviews.

The work developed in this dissertation is an attempt at validating aspects of Appraisal Theory in a contrastive manner (i.e. comparing English and Spanish), and at providing a cross-linguistic characterisation of this new review genre in terms of Appraisal features.

The research questions posed and investigated in this study are the following ones:

- (5) Is it possible to test the validity of aspects of Appraisal Theory empirically through corpus annotation? If so, are there features which need redefinition or reformulation to ensure high-quality bilingual (English-Spanish) annotation?
- (6) Are there language-independent Appraisal features which characterise mobile application reviews in English and Spanish? If so, which are these features?
- (7) Are there language-specific Appraisal features of mobile application reviews written in English and Spanish?
- (8) Are there significant differences in the use of Appraisal features depending on the type of item reviews and depending on the polarity features of the review?

In order to address these questions, this study was structured in two main parts: the empirical validation of Appraisal categories and the annotation of the larger corpus.

### **Empirical validation of Appraisal categories**

In order to validate empirically these categories as well as more delicate subcategories, three annotation experiments were designed, following the methodology proposed by Hovy and Lavid (2010), and Lavid (2012). These experiments were carried out on a training corpus of 50 bilingual (25 in English and 25 in Spanish) mobile application reviews posted on the application store for Android operating systems called Play Store addressing mobile applications, games, books, films and music.

The first annotation experiment addressed the identification of markables in the reviews. Coders were asked to indicate the number of spans that were supposed to involve any Appraisal meaning. The results showed a very high agreement in the length of the spans selected, but some disagreement occurred in the number of spans marked, which were mainly related to long phrases and sentences implying an evaluative meaning beyond the mere description of actions. With respect to coarse categories, Graduation appeared as the most conflictive category, which points to an unclear difference between intensification or additional description and values. Subcategories in Attitude (Affect, Judgement and Appreciation) showed the highest intersections, showing that even though Attitude is a category that can be clearly identified, subtypes within this category are not so noticeably different. This is closely related to the use of personifications (assignation of personal qualities to objects). Consequently, minor categories inside Appreciation,

more specifically, Valuation and Reaction, showed the highest variation between coders, where the use of terms other than adjectives made it more difficult to distinguish objects' qualities (Valuation) from objects as sources of others' qualities or emotions (Reaction).

The second experiment required the coders to label each of the spans previously marked with one tag from the core tagset (Attitude, Engagement, Graduation and their major subtypes Affect, Judgement, Appreciation, Expansion, Contraction, Force and Focus). A higher degree of agreement was obtained than in the first annotation experiment. Most disagreement was found in Graduation, followed by Engagement and, finally, Attitude. Graduation types are used to modify both Attitude and Engagement, and this may lead to some confusion regarding the actual Graduation meaning of intensification or softening and the Attitude and Engagement meanings to which those changes are directed.

The third annotation experiment addressed specific tags from the extended tagset. Moderate agreement was obtained when annotating specific subtypes of Attitude categories, where the confusion between Valuation and Reaction was especially noticeable. These categories involve the value and worthiness of the object reviewed, or a quality involving an impact on the user, respectively. The results for Engagement and Graduation showed an almost perfect agreement, and once these tags were selected, more delicate tags could also be distinguished by coders.

After these studies, all the categories annotated were validated, since they all showed a moderate and high agreement degree, and could be further used in the annotation of the larger corpus.

### **Annotation of the larger corpus**

The larger corpus used in this study consists of a collection of 200 texts (100 in English, 100 in Spanish) on the same products as those used in the training corpus. The annotation was carried out in the annotation environment UAM Corpus Tool. Apart from overall results, three aspects were compared: language-specific differences, item differences (applications, games, books, films and music) and polarity differences (positive and negative texts).

The results of the annotations in the larger set indicate interesting tendencies in the distribution of Appraisal tags in the bilingual reviews. First mobile applications reviews have been shown to be especially rich in Attitude tags, followed by Engagement, while

Graduation tags occur much less frequently. This distribution reflects the communicative purpose of these texts, which is to present users' opinions on a given product. Therefore, the majority of the Appraisal tags are expressions of Attitude which assign a value to the item reviewed, or express someone's feelings related to that item. The need to engage other users in the comments is also reflected in the quite abundant use of Engagement tags in both the English and the Spanish reviews. Graduation tags, used to intensify or soften ideas, appear much less frequently in these reviews, indicating that users prefer other Appraisal strategies to convey their opinions on a given product.

With respect to the comparison between English and Spanish texts, the latter use Attitude resources in a higher degree than the former, while this compensates its texts with higher Engagement and Graduation elements. Thus, Spanish texts draw on feelings and qualities much more frequently than English ones. The English reviews modulate their voice inside the text through Engagement as well as the words used by means of Graduation. Satisfaction meanings were preferred by the Spanish authors in a higher degree than the English ones, which focused on Happiness feelings. Regarding Engagement, the most noticeable aspects were the use of Epistemic and Pseudo-Questions in the case of English, and Deontic and Evidential markers in the case of Spanish. English modulate certainty more extensively and are also more sarcastic by using rhetoric questions in their writings while Spanish are much more direct using obligation resources and basing their opinion on empirical sources. Along the same lines, Counter elements are more profuse and varied in English, showing opposition and contrast, while Spanish writers are more direct by simply rejecting any other possibilities by means of Deny resources. Conclusions on Graduation include a higher use of Repetition in the Spanish texts, mostly realised by the repetition of exclamation marks. No other differences have been found regarding the use of Focus (either Sharpen or Soften), due to the limited representation in both languages, nor in Intensification or Quantification.

The linguistic realisations used in both languages presented in the most common spans in each of the three major axes, are very similar in English and in Spanish. Most of the linguistic items are actually shared as the most representative of each axis. Additionally, most of the most common linguistic items (typically, single words) were direct translations from one language into the other. They were broad generic wild terms such as *good* or *bad*, *better*, *awesome*, etc. Therefore, authors mostly use a very simplistic



language with highly recurrent words in common informal English and Spanish, instead of a field-specific terminology or jargon.

With respect to differences among products, some tendencies were observed that allowed the grouping of some products, although it is true that each product has its own preferences regarding the use of Appraisal items. The products could be arranged in two groups, one formed by applications and games, and a second one formed by books and films. Music shares some characteristics with both groups, but does also show its own proper qualities, as the other four types do. These subtle similarities may point to a distinction between traditionally reviewed items, such as books and films, and the new ones, applications and games, since they may also share some inherent qualities: books and films have a story, characters, a director or writer, etc. while applications and games deal with performance issues. Music does not attach to any other the two types completely and stands as a different group on its own.

Another comparison was established between positive texts, those with a rating of 4 or 5 stars, and negative texts, rated with 1 or 2 stars. As can be expected, positive texts were shown to include positive feelings like Happiness and Satisfaction, while Unhappiness and Dissatisfaction were found in negative texts, as well as a higher proportion of Deny elements. More interestingly, negative texts were also shown to include aspects which are typically criticised if they do not adjust to the author's expectations, but disregarded if they do, as was the case of Judgement elements. On the other hand, some other aspects are typically praised when they satisfy users, like the case of performance (Capacity).

To conclude, with the work reported in this dissertation, I hope to have shed light into a number of interesting issues related to Appraisal Theory and its use in the characterisation of mobile application reviews. First, the creation of a large bilingual annotated dataset containing Appraisal features will hopefully contribute to the advancement of knowledge in areas related to Sentiment Analysis in NLP and to the study of evaluation and interpersonal features of English and Spanish in the Corpus Linguistics community. It is also expected that the empirical validation of the annotation scheme proposed in this dissertation contributes to the refinement and reformulation of Appraisal Theory within the SFL community and encourages further applied work to other genres and other languages. Second, the bilingual annotation process of Spanish and

English mobile application reviews has provided an in-depth characterisation of this genre along Appraisal features, not only from the contrastive (English-Spanish) perspective, but also from the point of view of the type of item reviewed and the polarity assigned by the audience.

Future work will be focused on extending the current work to more specific lines of research in this area. Some of these lines include specific research on the realisation of Appraisal in long phrases and sentences, in order to find common validated features beyond readers' interpretations. Another interesting line of future research is the extension of the empirical validation of more delicate categories within Appraisal for which insufficient evidence was found in the current corpus. It would also be relevant to extend the current range of items reviewed to a wider range of products in order to find possible groups that share Appraisal features, thus confirming or diverging from the tendencies pointed out in this dissertation.

## Spanish summary

La evaluación, la opinión y la subjetividad son fenómenos relacionados que actualmente están recibiendo atención tanto por parte de la comunidad lingüística como de la computacional (Wiebe et al. 2004; Pang et al. 2002; Taboada 2009; Taboada and Grieve 2004; Whitelaw et al. 2005; Bloom, Garg y Argamon 2007; Bloom, Stein y Argamon 2007; Pang y Lee 2008, *inter alia*) bajo diferentes perspectivas y acercamientos.

Una de las teorías más influyentes que tratan el fenómeno de la evaluación y la subjetividad es la teoría de 'Appraisal' (Valoración<sup>7</sup>), una aproximación lingüística para 'explorar, describir y explicar la forma en que se usa el lenguaje para evaluar, indicar una postura, construir una imagen textual y administrar las relaciones y posicionamientos interpersonales' (ver página web de la teoría de la Valoración en <http://grammatics.com/appraisal/>). La teoría de la Valoración ha sido desarrollada por Martin y White en Australia (ver Martin 2000; Martin y White 2005; White 2003; White 2004, *inter alia*) y propone que las expresiones lingüísticas de significados evaluativos como la emoción, la actitud y la opinión se pueden divide en tres ejes diferentes: 'Attitude', 'Engagement' y 'Graduation' (Actitud, Intervención y Graduación).

---

<sup>7</sup> Todas las traducciones proporcionadas a los términos de las categorías se dan en traducción propia.

Como resultado del gran impacto de esta teoría, una cantidad considerable de investigación se ha dedicado a aplicar y expandir la teoría de diferentes maneras, centrándose principalmente en el área de Actitud (como por ejemplo, Macken-Horarik 2003; Painter 2003; Page 2003; White 2004), prestando menos atención a las áreas de Intervención y Graduación, que requieren un mayor estudio. La teoría de la Valoración se ha aplicado también a otros idiomas aparte del inglés y a distintos géneros (Sano 2011; Vian Jr. 2008; Pounds 2010; Becker 2009; Thomson, White y Kitley 2008; Taboada, Carretero y Hinnel 2014).

Sin embargo, hasta la fecha no existe un trabajo contrastivo a gran escala en la comunidad computacional o lingüística que haya intentado validar los diferentes rasgos de la teoría de la Valoración empíricamente a través de la anotación de corpus, e investigar su aplicación a un género relativamente nuevo como es el de las críticas de aplicaciones móviles en tiendas de aplicaciones.

El trabajo desarrollado en esta tesis es un intento de validar un conjunto de categorías de la teoría de la Valoración de manera contrastiva (comparando inglés y español), y proporcionar una caracterización de este nuevo género de críticas de acuerdo con los rasgos de la Valoración en ambas lenguas.

Las preguntas de investigación presentadas e investigadas en este estudio son las siguientes:

- (1) ¿Es posible probar la validez de varios aspectos de la teoría de la Valoración empíricamente a través de la anotación de corpus? Si es así, ¿hay rasgos que necesiten redefinirse o reformularse para garantizar una anotación bilingüe (inglés-español de calidad)?
- (2) ¿Existen rasgos de la Valoración independientes del idioma que caractericen las críticas en inglés y en español? Si es así, ¿cuáles son estos rasgos?
- (3) ¿Hay rasgos de la Valoración determinados por el idioma en las críticas de aplicaciones móviles escritas en inglés y en español?
- (4) ¿Hay diferencias estadísticamente significativas en el uso de los rasgos de la Valoración que dependan del tipo de producto comentado o que dependan de la polaridad de la crítica?

Para responder a estas preguntas, este estudio se ha estructurado en dos partes: la validación empírica de las categorías de la Valoración y la anotación de un corpus más extenso.

### **Validación empírica de las categorías de la Valoración**

Para validar empíricamente estas categorías además de otras categorías más específicas, se diseñaron tres experimentos de anotación, siguiendo la metodología propuesta por Hovy y Lavid (2010) y Lavid (2012). Estos experimentos se llevaron a cabo sobre un corpus de entrenamiento de 50 críticas (25 en inglés y 25 en español) de aplicaciones de móvil publicadas en la tienda de aplicaciones para el sistema operativo Android llamada Play Store, sobre aplicaciones, juegos, libros, películas y música para móviles.

El primer experimento de anotación se centró en la identificación de realizaciones de Valoración en las críticas. Los anotadores tenían que indicar el número de realizaciones que creían que tenían significados de Valoración. Los resultados mostraron un nivel de acuerdo muy alto en la longitud de las ocurrencias seleccionadas, aunque se vio algo más de desacuerdo en el número total de éstas, debido principalmente a la longitud de los sintagmas y oraciones que conllevaban significados evaluativos más allá de la mera descripción de acciones. Con respecto a categorías más generales, la Graduación se mostró como la categoría más conflictiva, lo cual apunta a la falta de claridad en la diferencia entre la intensificación y la descripción y valoración adicionales. Las categorías dentro de Actitud ('Affect', 'Judgement' y 'Appreciation', Afecto, Juicio y Apreciación) mostraron los mayores solapamientos, de modo que aunque Actitud es una categoría que se puede identificar claramente, los subtipos dentro de esta categoría no se distinguen tan fácilmente. Esto está estrechamente relacionado con el uso de personificaciones (la asignación de cualidades personales a objetos). Como consecuencia, las categorías más específicas dentro de Apreciación, más concretamente, 'Valuation' y 'Reaction' (Tasación y Reacción), mostraron los valores más altos de desacuerdo entre los anotadores, haciendo el uso de términos diferentes a adjetivos más difícil distinguir las cualidades de los objetos (Tasación) de los objetos como fuentes de las cualidades y emociones de otros (Reacción).

En el segundo experimento se pedía a los anotadores que asignaran una categoría a cada una de las realizaciones señaladas previamente con una etiqueta del listado de etiquetas central (Actitud, Intervención, Graduación y sus categorías principales, Afecto,

Juicio, Apreciación, Expansión, Contracción, Fuerza y Foco). El nivel de acuerdo que se obtuvo fue superior al del primer experimento de anotación. La mayor parte del desacuerdo se encontró en Graduación, seguido de Intervención y, por último, Actitud. Los tipos de Graduación se usan para modificar tanto la Actitud como la Intervención, y esto puede provocar cierta confusión entre el significado real de Graduación de intensificación o suavizado y los significados a los que esto va dirigido.

El tercer experimento de anotación consistió en la asignación de etiquetas del listado ampliado. Se obtuvo un nivel de acuerdo moderado al anotar algunos subtipos concretos de Actitud, donde la confusión entre Tasación y Reacción fue especialmente notable. Estas categorías cubren el valor y valía del objeto criticado, o una cualidad que implique un efecto en el usuario, respectivamente. Por su parte, los resultados para Intervención y Graduación mostraron un acuerdo casi perfecto.

Tras estos estudios de fiabilidad, se validaron todas las categorías anotadas, ya que todas mostraron un nivel de acuerdo moderado o alto, y se pudieron usar posteriormente para anotar el corpus más extenso.

### **Anotación del corpus más extenso**

El corpus más extenso que se ha usado en este estudio consiste de una compilación de 200 textos (100 en inglés y 100 en español) sobre los mismos productos usados en el corpus de entrenamiento. La anotación se llevó a cabo usando el entorno de anotación UAM Corpus Tool. Además de obtener resultados globales, se compararon tres aspectos: diferencias en cuanto al idioma, diferencias en cuanto al producto criticado (aplicaciones, juegos, libros, películas y música) y diferencias en cuanto a la polaridad (textos positivos y negativos).

Los resultados de las anotaciones de este corpus muestran tendencias interesantes en la distribución de la Valoración en las críticas bilingües. En primer lugar, las críticas de aplicaciones móviles han mostrado ser especialmente ricas en realizaciones de Actitud, seguidas de Intervención, mientras que la Graduación ocurría con mucha menos frecuencia. Esta distribución refleja el propósito comunicativo de estos textos: presentar las opiniones de los usuarios sobre un producto dado. Por tanto, la mayor parte de la Valoración se trata de expresiones de Actitud que asignan un valor al objeto criticado, a expresan los sentimientos de alguien en relación a ese objeto. La necesidad de conectar con otros usuarios en los comentarios se ve también en abundante uso de realizaciones de

Intervención tanto en las críticas en inglés como en las que estaban en español. La Graduación, usada para intensificar o suavizar ideas, aparece con mucha menos frecuencia en estas críticas, lo que indica que los usuarios prefieren otras estrategias de Valoración para transmitir sus opiniones sobre un producto.

Con respecto a la comparación entre los textos en inglés y en español, los últimos utilizan recursos de Actitud en mayor grado que los primeros, mientras que los textos ingleses utilizan en mayor medida elementos de Intervención y Graduación. De este modo, los textos en español se apoyan en los sentimientos y cualidades con mayor frecuencia que los ingleses. Las críticas en inglés modulan su voz dentro del texto a través de la Intervención, además de otras palabras utilizadas mediante la Graduación. Los significados de ‘Satisfaction’ (Satisfacción) se prefieren en los textos en español en mayor medida que en los ingleses, los cuales se centran más en sentimientos de ‘Happiness’ (Felicidad). Respecto a la Intervención, los aspectos más notables fueron el uso de las categorías ‘Epistemic’ y ‘Pseudo-Questions’ (Epistémico y Pseudo-Preguntas) en el caso del inglés, y ‘Deontic’ y ‘Evidential’ (Deóntico y Evidencial) en el caso del español. En los textos en inglés se modula la certeza en mayor grado y son también más sarcásticos mediante el uso de preguntas retóricas, mientras que los textos en español son más directos al usar recursos de obligación y basar sus opiniones en fuentes empíricas. En la misma línea, los elementos de ‘Counter’ (Oposición), que muestran oposición y contraste, son más profusos y variados en inglés, mientras que los autores españoles son más directos rechazando directamente otras posibilidades mediante el uso de ‘Deny’ (Negación). Las conclusiones sobre el uso de la Graduación incluyen un mayor uso de ‘Repetition’ (Repetición) en los textos españoles, principalmente realizados mediante la repetición de signos exclamativos. No se encontraron diferencias respecto al uso de Foco —ni para ‘Sharpen’ (Acentuación) ni para ‘Soften’ (Suavizado)— debido a la limitada representación que tuvieron estas categorías en ambos idiomas, así como ni para ‘Intensification’ o ‘Quantification’ (Intensificación y Cuantificación).

Las realizaciones lingüísticas que se usaron en ambos idiomas presentaron una gran parte de ejemplos en común para los tres ejes principales. La mayoría de los términos lingüísticos eran compartidos como los más utilizados. Además, la mayoría de estos términos más frecuentes eran traducciones directas de un idioma al otro. Se trata de palabras comodín como *bueno* y *malo*, *mejor*, *genial*, etc. Por tanto, se puede decir que

los autores utilizan un lenguaje muy simple con un gran uso de palabras recurrentes del inglés y español informales, en lugar de una terminología o jerga específicas del campo.

Con respecto a las diferencias entre productos, se observaron algunas tendencias que permitieron agrupar algunos productos, aunque es cierto que cada producto tiene sus preferencias respecto al uso de términos de Valoración. Se podían establecer dos grupos: uno formado por las aplicaciones y los juegos, y otro por los libros y las películas. La música comparte algunas características con ambos grupos, pero también tiene las suyas propias, tal y como hacen los otros cuatro grupos. Estas pequeñas similitudes pueden apuntar a una posible distinción entre los productos criticados tradicionalmente (los libros y las películas) y los nuevos (las aplicaciones y los juegos), puesto que pueden compartir algunas características inherentes: los libros y las películas tienen una historia, personajes, un director o escritor, etc. mientras que de las aplicaciones y los juegos se suele señalar su rendimiento. La música no encaja con ninguno de los dos tipos por completo, y se presenta como grupo aparte.

Una tercera comparación se estableció entre los textos positivos, aquellos con una valoración de 4 ó 5 estrellas, y los negativos, con 1 ó 2 estrellas. Como cabe esperar, los textos positivos mostraron en mayor medida sentimientos como Felicidad o Satisfacción, mientras que se encontró más Infelicidad e Insatisfacción ('Unhappiness' y 'Dissatisfaction') en los negativos, así como una mayor proporción de elementos de Negación. Curiosamente, los textos negativos también incluyeron aspectos que se critican habitualmente cuando no cumplen con las expectativas del autor, pero que se pasan por alto si las cumplen, como es el caso de los elementos de Juicio. Por otro lado, algunos otros aspectos se alaban habitualmente cuando satisfacen a los usuarios, como es el caso del rendimiento o ejecución, representado por la categoría 'Capacity' (Capacidad).

Para concluir, con el trabajo presentado en esta tesis, espero haber arrojado algo de luz sobre una serie de cuestiones relacionadas con la teoría de la Valoración y su uso en la caracterización de las críticas de aplicaciones móviles. En primer lugar, es de esperar que la creación de un conjunto extenso de datos bilingües anotados con rasgos de Valoración contribuya al avance del conocimiento en áreas relacionadas con el Análisis de Sentimiento en el Procesamiento del Lenguaje Natural (PLN), por una parte, y al estudio de la evaluación y los rasgos interpersonales del inglés y el español en el ámbito de la Lingüística de Corpus, por otro. También es de esperar que la validación empírica

del esquema de anotación propuesta en esta tesis contribuya a la mejora y reformulación de la teoría de la Valoración dentro de la comunidad de Lingüística Sistémico-Funcional y anime a realizar trabajos posteriores sobre otros géneros e idiomas. En segundo lugar, el proceso de anotación bilingüe de críticas de aplicaciones móviles en inglés y en español ha proporcionado una caracterización en profundidad de este género de acuerdo con los rasgos de la Valoración, no sólo desde una perspectiva contrastiva (inglés-español), sino también desde el punto de vista del producto criticado y la polaridad asignada por el usuario.

En cuanto a las líneas de investigación futura que se abren con este trabajo, son de destacar las siguientes: en primer lugar, una línea relevante es el estudio de las realizaciones lingüísticas de la Valoración mediante sintagmas y oraciones más largos, con el fin de encontrar rasgos validados comunes más allá de la interpretación del lector. Otra línea de investigación interesante es la extensión de la validación empírica de las categorías más específicas dentro de la Valoración que han tenido una representación insuficiente en este trabajo. También sería relevante extender la variedad de productos criticados con un abanico más amplio para encontrar posibles agrupaciones que compartan rasgos de Valoración, confirmando así o distanciándose de las tendencias señaladas en esta tesis.



## Appendix C. QUANTIFICATION OF TOTAL COMMENTS EXTRACTED

	English		Spanish	
<b>Clean Master</b>	5 stars: 3182	2 stars: 86	5 stars: 2805	2 stars: 116
	4 stars: 356	1 star: 158	4 stars: 547	1 star: 212
	<i>Total: 3782 comments</i>		<i>Total: 3680 comments</i>	
<b>Duolingo</b>	5 stars: 3506	2 stars: 18	5 stars: 3279	2 stars: 22
	4 stars: 647	1 star: 41	4 stars: 745	1 star: 23
	<i>Total: 4212 comments</i>		<i>Total: 4069 comments</i>	
<b>Instagram</b>	5 stars: 2040	2 stars: 383	5 stars: 2757	2 stars: 166
	4 stars: 707	1 star: 642	4 stars: 577	1 star: 322
	<i>Total: 3772 comments</i>		<i>Total: 3822 comments</i>	
<b>WhatsApp</b>	5 stars: 2210	2 stars: 322	5 stars: 1488	2 stars: 407
	4 stars: 775	1 star: 648	4 stars: 811	1 star: 948
	<i>Total: 3955 comments</i>		<i>Total: 3654 comments</i>	
<b>Total</b>	<i>15721 English comments</i>		<i>15225 Spanish comments</i>	

**Table C21. Total number of application comments extracted**

	English		Spanish	
<b>Angry Birds</b>	5 stars: 2767	2 stars: 137	5 stars: 2864	2 stars: 126
	4 stars: 476	1 star: 461	4 stars: 523	1 star: 321
	<i>Total: 3841 comments</i>		<i>Total: 3834 comments</i>	
<b>Candy Crush</b>	5 stars: 2478	2 stars: 175	5 stars: 2018	2 stars: 251
	4 stars: 648	1 star: 335	4 stars: 725	1 star: 921
	<i>Total: 3636 comments</i>		<i>Total: 3915 comments</i>	
<b>Clash of Clans</b>	5 stars: 1885	2 stars: 270	5 stars: 1927	2 stars: 278
	4 stars: 654	1 star: 852	4 stars: 963	1 star: 719
	<i>Total: 3661 comments</i>		<i>Total: 3887 comments</i>	
<b>Pou</b>	5 stars: 3224	2 stars: 116	5 stars: 2149	2 stars: 211
	4 stars: 511	1 star: 299	4 stars: 909	1 star: 423
	<i>Total: 4150 comments</i>		<i>Total: 3692 comments</i>	
<b>Total</b>	<i>15288 English comments</i>		<i>15328 Spanish comments</i>	

**Table C22. Total number of game comments extracted**

	English		Spanish	
<b><i>All the Light We Cannot See</i></b>	5 stars: 147	2 stars: 3	5 stars: 22	2 stars: 0
	4 stars: 30	1 star: 1	4 stars: 5	1 star: 1
	<i>Total: 181 comments</i>		<i>Total: 28 comments</i>	
<b><i>Fifty Shades of Grey</i></b>	5 stars: 3173	2 stars: 65	5 stars: 1639	2 stars: 25
	4 stars: 295	1 star: 426	4 stars: 134	1 star: 122
	<i>Total: 3959 comments</i>		<i>Total: 1920 comments</i>	
<b><i>The Girl on the Train</i></b>	5 stars: 456	2 stars: 15	5 stars: 162	2 stars: 17
	4 stars: 133	1 star: 26	4 stars: 60	1 star: 21
	<i>Total: 630 comments</i>		<i>Total: 260 comments</i>	
<b><i>The Life-Changing Magic of Tidying Up</i></b>	5 stars: 93	2 stars: 8	5 stars: 11	2 stars: 1
	4 stars: 32	1 star: 6	4 stars: 3	1 star: 0
	<i>Total: 139 comments</i>		<i>Total: 15 comments</i>	
<b><i>Total</i></b>	<i>4909 English comments</i>		<i>2223 Spanish comments</i>	

**Table C23. Total number of book comments extracted**

	English		Spanish	
<b><i>Avatar</i></b>	5 stars: 695	2 stars: 37	5 stars: 179	2 stars: 8
	4 stars: 60	1 star: 64	4 stars: 19	1 star: 65
	<i>Total: 856 comments</i>		<i>Total: 271 comments</i>	
<b><i>Gravity</i></b>	5 stars: 817	2 stars: 153	5 stars: 23	2 stars: 5
	4 stars: 230	1 star: 304	4 stars: 14	1 star: 10
	<i>Total: 1504 comments</i>		<i>Total: 52 comments</i>	
<b><i>Frozen</i></b>	5 stars: 3048	2 stars: 109	5 stars: 566	2 stars: 15
	4 stars: 253	1 star: 508	4 stars: 36	1 star: 49
	<i>Total: 3918 comments</i>		<i>Total: 666 comments</i>	
<b><i>Sherlock Holmes</i></b>	5 stars: 255	2 stars: 8	5 stars: 235	2 stars: 10
	4 stars: 50	1 star: 20	4 stars: 48	1 star: 98
	<i>Total: 333 comments</i>		<i>Total: 391 comments</i>	
<b><i>The Wolf of Wall Street</i></b>	5 stars: 924	2 stars: 40	5 stars: 160	2 stars: 10
	4 stars: 108	1 star: 110	4 stars: 21	1 star: 24
	<i>Total: 1182 comments</i>		<i>Total: 215 comments</i>	
<b><i>Total</i></b>	<i>7793 English comments</i>		<i>1595 Spanish comments</i>	

**Table C24. Total number of film comments extracted**

	English		Spanish	
<b>AC/DC</b>	5 stars: 547	2 stars: 1	5 stars: 214	2 stars: 0
	4 stars: 27	1 star: 9	4 stars: 13	1 star: 5
	<i>Total: 548 comments</i>		<i>Total: 232 comments</i>	
<b>Adele</b>	5 stars: 3429	2 stars: 69	5 stars: 1871	2 stars: 25
	4 stars: 211	1 star: 244	4 stars: 107	1 star: 113
	<i>Total: 3953 comments</i>		<i>Total: 2116 comments</i>	
<b>Eminem</b>	5 stars: 616	2 stars: 2	5 stars: 74	2 stars: 0
	4 stars: 35	1 star: 24	4 stars: 2	1 star: 3
	<i>Total: 677 comments</i>		<i>Total: 79 comments</i>	
<b>Michael Jackson</b>	5 stars: 608	2 stars: 3	5 stars: 385	2 stars: 2
	4 stars: 17	1 star: 15	4 stars: 9	1 star: 9
	<i>Total: 643 comments</i>		<i>Total: 405 comments</i>	
<b>Pitbull</b>	5 stars: 126	2 stars: 4	5 stars: 68	2 stars: 4
	4 stars: 9	1 star: 16	4 stars: 6	1 star: 23
	<i>Total: 155 comments</i>		<i>Total: 101 comments</i>	
<b>Total</b>	<i>5976 English comments</i>		<i>2933 Spanish comments</i>	

Table C25. Total number of music comments extracted



## Appendix D. ANNOTATION GUIDELINES

(Adapted from Martin and White 2005)

### 1. ATTITUDE

**Definition:** Attitude is concerned with our feelings, including emotional reactions, judgements of behaviour and evaluation of things.

#### 1.1. AFFECT

**Definition:** Emotional reactions and feelings, either positive or negative.

##### 1.1.1. HAPPINESS

**Definition:** Feelings of cheer and affection.

**Examples:** *chuckle, laugh, rejoice, cheerful, jubilant, shake hands, hug, embrace, be fond of, love, adore.* *sonreír, reír, regocijarse, alegre, jubiloso, estrechar la mano, abrazar, amar, adorar, ser aficionado a, idolatrar, venerar.*

##### 1.1.2. UNHAPPINESS

**Definition:** Feelings of misery and antipathy.

**Examples:** *whimper, cry, wail, sad, miserable, rubbish, abuse, revile, dislike, hate, abhor.* *quejarse, llorar, gemir, triste, miserable, insultar, criticar, faltar a, detester, desagradar, aborrecer, odiar.*

##### 1.1.3. SECURITY

**Definition:** Feelings of confidence and trust.

**Examples:** *delegate, commit, entrust, comfortable with, confident in/about, trusting.* *confiar, fiarse de, seguro, confiado, suelto, firme, con soltura.*

##### 1.1.4. INSECURITY

**Definition:** Feelings of disquiet and surprise.

**Examples:** *restless, twitching, shaking, cry out, faint, uneasy, anxious, freaked out, started, jolted, staggered* *inquieto, alterado, nervioso, tenso, crispado, dudoso, temeroso.*

### 1.1.5. SATISFACTION

**Definition:** Feelings of interest and pleasure.

**Examples:** *attentive, busy, industrious, pat on the back, compliment, reward, involved, absorbed, engrossed, satisfied, pleased, charmed, chuffed, thrilled* *interesado, ilusionado, encantado, esperanzado, con expectativas, implicado, absorto, dedicado, empeñado, satisfecho, encantado, impresionado.*

### 1.1.6. DISSATISFACTION

**Definition:** Feelings of ennui and displeasure.

**Examples:** *fidget, yawn, tune out, caution, scold, castigate, flat, jaded, stale, , bored with, angry, sick of, furious, fed up with* *bostezar, adormilarse, desconectar, aburrirse, aburrido, harto, hastiado, quemado, desilusionado, desencantado, abatido.*

## 1.2. JUDGEMENT

**Definition:** Assessment of behaviour according to normative principles.

### 1.2.1. NORMALITY

**Definition:** How (un)usual someone or something is

**Examples:** *lucky, fortunate, charmed, normal, familiar, usual, natural, cool, stable, predictable, in, fashionable, avant garde, celebrated, unsung, unlucky, hapless, star-crossed, odd, peculiar, eccentric, erratic, unpredictable, dated, retrograde* *suertudo, afortunado, con estrella, normal, familiar, usual, natural, frío, estable, predecible, de moda, de vanguardia, no reconocido, gafe, raro, excéntrico, impredecible, anticuado*

### 1.2.2. CAPACITY

**Definition:** How (in)capable someone or something is

**Examples:** *powerful, vigorous, robust, adult, mature, experienced, witty, humorous, insightful, clever, gifted, sensible, expert, competent, successful, productive.* *Poderoso, vigoroso, adulto, maduro, experimentado, ingenioso, chistoso, perspicaz, inteligente, dotado, sensible, experto, competente.*

<i>weak, immature, childish, dull, dreary, slow, stupid, thick, naïve, inexpert, foolish, illiterate, uneducated, ignorant, incompetent, unsuccessful</i>	<i>débil, inmaduro, infantil, aburrido, lúgubre, lento, estúpido, pensar, ingenuo, inexperto, insensato, analfabeto, ignorante, incompetente, sin éxito</i>
---	---

### 1.2.3. TENACITY

**Definition:** How (ir)resolute someone or something is

**Examples:** *brave, heroic, cautious, patient, valiente, heroic, precavido,  
careful, meticulous, tireless, prudente, paciente, constante,  
persevering, resolute, flexible tenaz, perseverante, infatigable,  
flexible.*

<i>timid, impatient, capricious, stubborn, obstinate.</i>	<i>tímido, cobarde, impaciente, voluble, terco, obstinado.</i>
---	--

### 1.2.4. VERACITY

**Definition:** How truthful someone or something is

**Examples:** *truthful, honest, credible, frank, honrado, creíble, sincero, franco,  
discrete, tactful. directo, discreto, diplomático*

<i>dishonest, deceitful, lying, deceptive, manipulative, blunt.</i>	<i>deshonesto, mentiroso, engañoso, manipulador, taimado.</i>
---	---

### 1.2.5. PROPRIETY

**Definition:** How ethical someone or something is

**Examples:** *moral, ethical, fair, sensitive, kind, moral, ético, justo, sensible,  
caring, modest, humble, polite, amable, cariñoso, modesto,  
respectful, generous, charitable. humilde, educado, respetuoso,  
generoso, caritativo*

<i>immoral, evil, corrupt, unfair, insensitive, mean, cruel, arrogant, rude, selfish, greedy, avaricious.</i>	<i>immoral, malvado, corrupto, injusto, insensible, cruel, arrogante, grosero, egoísta, codicioso, avaricioso</i>
---	---

## 1.3. APPRECIATION

**Definition:** Evaluation and valuation of things.

### 1.3.1. REACTION

**Definition:** Effect, impact and appearance.

**Examples:** *arresting, captivating, engaging, fascinating, exciting, moving, dramatic, intense, remarkable, sensational, lovely, beautiful, appealing, enchanting.*

*atractivo, fascinante, emocionante, excitante, intenso, notable, sensacional, encantador, hermoso, atractivo.*

*dull, boring, tedious, dry, uninviting, flat, predictable, monotonous, unremarkable, ugly, grotesque, repulsive*

*aburrido, tedioso, seco, poco atractivo, plano, predecible, monótono, nada especial, feo, grotesco, repulsive.*

### 1.3.2. COMPOSITION

**Definition:** Balance and complexity.

**Examples:** *balanced, harmonious, unified, symmetrical, proportioned, consistent, logical, simple, pure, elegant, lucid, clear, precise, intricate, rich, detailed*

*equilibrado, armonioso, unificado, simétrico, proporcional, coherente, lógico, simple, puro, elegante, lúcido, claro, preciso, intrincado, rico, detallado.*

*unbalanced, discordant, irregular, uneven, contradictory, disorganised, distorted, extravagant, ornate, unclear, simplistic*

*desequilibrado, discordante, irregular, contradictorio, desorganizado, distorsionado, extravagante, adornado, poco claro, simplista.*

### 1.3.3. VALUATION

**Definition:** Value and worthiness.

**Examples:** *profound, innovative, original, creative, inimitable, exceptional, unique, valuable, priceless, worthwhile, helpful, effective*

*profundo, innovador, original, creativo, inimitable, excepcional, único, valioso, inestimable, útil, eficaz.*

*insignificant, conventional, worthless, ineffective, useless*

*insignificante, convencional, inútil, ineficaz.*



## 2. ENGAGEMENT

**Definition:** Implication of other possibilities and voices than the speakers’.

### 2.1. EXPANSION

**Definition:** Author’s position is one inside a range of possible options or an external source provided for a given opinion.

#### 2.1.1. EPISTEMIC

**Definition:** Degree of certainty

<b>Examples:</b>	<i>may, might, perhaps, probably, definitely</i>	<i>poder, quizás, probablemente, definitivamente</i>
	<i>I think, I believe, I’m convinced</i>	<i>pienso, creo, estoy convencido</i>
	<i>they may be late</i>	<i>puede que lleguen tarde</i>
	<i>it’s likely that</i>	<i>es probable que...</i>
	<i>in my opinion, in my point of view</i>	<i>en mi opinión, desde mi punto de vista</i>

#### 2.1.2. EVIDENTIAL

**Definition:** Empirical evidence

<b>Examples:</b>	<i>seem, suggest, apparently</i>	<i>parecer, sugerir, aparentemente</i>
	<i>it seems that...</i>	<i>parece que...</i>
	<i>this suggests that...</i>	<i>esto sugiere que...</i>

#### 2.1.3. PSEUDO-QUESTION

**Definition:** Questions with unintended answer

<b>Examples:</b>	<i>what were they thinking?</i>	<i>¿en qué estaban pensando?</i>
	<i>how can they be so stupid?</i>	<i>¿cómo pueden ser tan tontos?</i>

#### 2.1.4. DEONTIC

**Definition:** Obligation, control and compliance

<b>Examples:</b>	<i>must, should, have to</i>	<i>deber, tener que</i>
------------------	------------------------------	-------------------------

#### 2.1.5. ACKNOWLEDGE

**Definition:** External source provided for a given opinion with no overt indication as to where the authorial voice stands

<b>Examples:</b>	<i>say, declare, believe, think</i> <i>X said that...</i> <i>according to X...</i> <i>in X's view...</i>	<i>decir, declarar, creer, pensar</i> <i>X dijo que...</i> <i>según X...</i> <i>en opinión de X...</i>
------------------	---	---

### 2.1.6. DISTANCE

**Definition:** External source provided for a given opinion with an explicit distancing of the authorial

<b>Examples:</b>	<i>claim, rumour</i> <i>X claims that...</i> <i>it's rumoured that...</i>	<i>afirmar, rumorear</i> <i>X afirmar que...</i> <i>se rumorea que...</i>
------------------	---	---

## 2.2. CONTRACTION

**Definition:** Author positions against a contrary position or limits the scope of possibilities.

### 2.2.1. DENY

**Definition:** Negation.

<b>Examples:</b>	<i>not, no, never</i>	<i>no, ningún, nunca</i>
------------------	-----------------------	--------------------------

### 2.2.2. COUNTER

**Definition:** Concession and counter expectation.

<b>Examples:</b>	<i>although, however, yet, but,</i> <i>surprisingly, even</i>	<i>aunque, sin embargo, aún así,</i> <i>pero, sorprendentemente, incluso,</i> <i>siquiera</i>
------------------	--	---

### 2.2.3. CONCUR

**Definition:** Overtly announce the addresser as agreeing with, or having the same knowledge as, some projected alternative.

<b>Examples:</b>	<i>naturally..., of course...,</i> <i>obviously..., certainly...</i>	<i>naturalmente..., por supuesto...,</i> <i>obviamente..., ciertamente...</i>
------------------	---	--

### 2.2.4. PRONOUNCE

**Definition:** It involved authorial emphases or explicit authorial interventions or interpolations.

**Examples:** *I conted..., the truth is..., there can be no doubt that..., the facts of the matter are that..., we can only conclude that..., you must agree that...*      *yo sostengo que..., la verdad es..., no puede haber duda de que... los hechos son que..., solo podemos concluir que..., debes estar de acuerdo en que...*

#### 2.2.5. ENDORSE

**Definition:** Propositions sourced to external sources are construed by the authorial voice as correct, valid and undeniable.

**Examples:** *show, prove, demonstrate*      *mostrar, probar, demostrar*  
*X has demonstrated that...*      *X ha demostrado que...*  
*as X has shown...*      *como X ha mostrado...*

### 3. GRADUATION

**Definition:** Grading phenomena whereby feelings are amplified or softened and categories blurred or sharpened.

#### 3.1. FOCUS

**Definition:** Degree of prototypicality, i.e. reference to the degree to which they match some supposed core or exemplary instance of a semantic category.

##### 3.1.1. SHARPEN

**Definition:** Maximisation of the degree of prototypicality.

**Examples:** *a real father*      *un verdadero padre*  
*a true friend*      *un amigo de verdad*  
*a genuine hero*      *un héroe genuino*

##### 3.1.2. SOFTEN

**Definition:** Lessening of the degree of prototypicality.

**Examples:** *Is he your boyfriend? –Kind of.*      *¿Es tu novio? –Algo así.*  
*it was an apology of sorts*      *era una especie de disculpa*

#### 3.2. FORCE

**Definition:** Degree of intensity or amount.

### 3.2.1. ISOLATION

**Definition:** Scaling of intensity realised by an individual lexical item.

**Examples:** *very, so, too, quite, slightly, extremely*      *muy, tan, demasiado, bastante, ligeramente, extremadamente*  
*yes!*      *¡sí!*

### 3.2.2. INFUSION

**Definition:** Scaling of intensity realised by one aspect of the meaning of a single term.

**Examples:** *(good) better, best*      *(bueno), mejor, el mejor*  
*rarely, occasionally, sometimes, often, always*      *rara vez, ocasionalmente, a veces, a menudo, siempre*  
*it's VERY important*      *es MUY importante*

### 3.2.3. REPETITION

**Definition:** Repetition of the same lexical item or part of it.

**Examples:** *it's hot hot hot*      *está caliente caliente caliente*  
*we laughed and laughed and laughed*      *reímos y reímos y reímos*  
*yeeeeeeeees*      *síiiiiiiii*  
*yes!!!!!*      *¡¡¡¡sí!!!!!*

### 3.2.4. NUMBER

**Definition:** Numerical quantifiers.

**Examples:** *few, lots of, many, heaps of*      *pocos, montones, muchos, pilas de*

### 3.2.5. MASS

**Definition:** Extension of an item's presence in space.

**Examples:** *tiny, small, large, huge, gigantic*      *diminuto, pequeño, grande, enorme, gigante*

### 3.2.6. EXTENT

**Definition:** Proximity and distribution in time and space.

<b>Examples:</b>	<u>recent</u> arrival	<u>reciente</u> llegada
	<u>ancient</u> betrayal	<u>antigua</u> traición
	<u>long-lasting</u> hostility	hostilidad <u>duradera</u>
	<u>short</u> battle	batalla <u>corta</u>
	<u>wide-spreading</u> hostility	hostilidad <u>de amplia difusión</u>
	<u>narrowly-based</u> support	soporte <u>estrecho</u>



## Appendix E. TRAINING CORPUS

**Language:** en (English)  
es (Spanish)

**Application's name:** Clean Master (com.cleanmaster.mguard)  
Instagram com.instagram.android)  
Candy Crush (com.king.candycrushsaga)  
Angry Birds (com.rovio.angrybirds),  
All the Light We Cannot See  
Fifty shades of grey  
The Girl on the Train  
Avatar  
Frozen (Frozen 2013)  
Gravity  
The Wolf of Wall Street  
AC/DC (AC\_DC Back in Black)  
Adele (Adele25)

**Stars:** from 1 to 5.

**Review number:** from 1 to 200.

**Title example:**

<i>en-</i>	<i>com.cleanmaster.mguard-</i>	<i>1-</i>	<i>T0</i>
language	application's name	stars	review number

### 1. APPLICATIONS

#### app - es - whatsapp - 5 - T1

Nueva opción de recomendación

El WhatsApp es súper bueno pero quiten las notificaciones flotantes dejen las como antes y la nueva opción de la que les iba hablar era que en cada contacto aparezca al lado si esta en línea o no como tipo Skype que aparece al lado del contacto una bolita verde si esta conectado, un círculo verde si se desconectó hace poco y un círculo gris si lleva varios dias desconectado. Espero que sea escuchado; muchas gracias

#### app - es - whatsapp - 4 - T2

Bien

Hasta el momento todo marcha bien, como todo programa nunca sera perfecto con el paso del tiempo mejora y tienen que ver los detalles que como usuarios tenemos para de esta forma hacerla cada vez mejor. Gracias por permitirme comunicar con mis seres queridos.

#### app - es - whatsapp - 2 - T3

Funciona muy mal.

Es muy inestable.... Estoy conectado cuando ni siquiera tengo encendido mi celular no llegan los mensajes a tiempo cuando me los envían antes no sucedía pero bueno.... Solucionen eso porque seguramente no soy el único a quién le sucede así qué tengan en cuenta mi humilde opinión y arreglen estos detalles.

**app - es - whatsapp - 1 - T4**

Vaya mierda de nueva actualización

No me gusta nada la nueva actualización de esta semana porque no aparece el luminoso del Led cuando entra un mensaje y ya no se pueden leer en la cortinilla sin que se vea que se chequean. Por favor, espero que se cambie porque así no hay quien se maneje con él. Actualizarlo Yaaaaaa

**app - es - whatsapp - 1 - T5**

Lo peor de lo peor

No he visto nunca una app como está con tantas actualizaciones y lo peor es que cada vez va peor pesa como un mastodonte y me ralentiza muchísimo mi teléfono está app no es lo que era ahora parece un ladrillo y cada vez que la actualizan que son muchísimas veces que encima no es para tener más características la engordan tanto que es un auténtico dinosaurio come recursos, nos estamos planteando cambiar a otra app que hay muchísimas mejores y más ligeritas arreglen esto o nos vamos

**app - en - whatsapp - 5 - T6**

Best Messaging App over the World

Good app, no need to explain. It is revolution in Messaging App. What's up app is one of my favourite app and many people like this. One issue with what's up app regarding calling which is fail. Need more improvement over calling features. PDF from is only supposed is another issue please allow docx format so easily access the what's up app.

**app - en - whatsapp - 4 - T7**

Needs improvement/more features

Great app overall. Love how you can chat from your computer- its super convenient. No ads, great performance, and app is very fast overall. However, this app could use some improvements. The number 1 thing that really bothers me is that the phone automatically saves any shared pictures/videos/MMS. I've tried looking how to disable this on the internet and it does not work... Also, being hopeful that ONE DAY video chatting exists in whatsapp. Other than that, I really enjoy this app!

**app - en - whatsapp - 2 - T8**

Annoying notification system

It would be good not to push every single message in conversation, just the first one. In other words if I did not read the 1 st one, I don't want to be bothered next 15 times. At the same time, this should be per conversation I.e. still push the message in new conversation.



**app - en - whatsapp - 1 - T9**

New update is rubbish.

New update sucks. Everytime I'm on a whatsapp call always says 'reconnecting' can never hear the person on the other end and neither can they hear me, didn't have this problem with the older version so no its not my phone or network. Absoloutely hate the new update, prefer the older one would have rated it 5 stars but due to the new update its a no from me.

**app - en - whatsapp - 1 - T10**

More useless features

Crap since last update, always showing the wrong message in the pop-up, keeps showing unread message count in the chat window of the person is talking to. Just what I needed. Another f\*ing banner on the screen that won't go away until I reply that tells me how man unread messages I have. Do you really think people are that stupid that they do not know they have messages while interacting with the person in the same chat window?

**2. GAMES**

**game - es - candycrushsaga - 5 - T11**

Me encanta

Es divertido y no te aburres lo que pasa que algunos niveles son muy dificiles y te enrritas yo me enrrite me lo desintale y como me aburria y con este juego me divertia pues me lo e vuelto s descargar si te aburres te lo recomiendo divierte mucho y esta chulisimo es unos de los juegos que mas me gustan seguir asi de bien

**game - es - candycrushsaga - 5 - T12**

Candy crush, un juego genial

Este juego es el típico de juntar de a tres, pero diganle quien no ha jugado este tipo de juegos alguna vez, es un éxito en mi familia y aún fuera de ella, cuando todos tenemos gustos diferentes, en fin, Yo lo recomiendo

**game - es - candycrushsaga - 4 - T13**

Me entretiene

Es un juego bastante divertido, que ha incluido muchas cosas, como nuevos Bosters, y es increíble que cada vez haya niveles nuevos, simplemente un juego que te pone a trabajar la mente, y eso es muy bueno.

**game - es - candycrushsaga - 2 - T14**

Mala actualización

Antes podía dejarlo en segundo plano, abrir otras apps o dejar el celular y luego abriría a mitad de mi juego. Ahora si cambio de app, pierdo el juego abierto. Odio eso. Voy a desinstalarlo y buscar una versión anterior.

**game - es - candycrushsaga - 1 - T15**

Harta!

Estoy harta! Llevo varios años jugando a este juego y ya es la segunda vez que me quita todas mis ayudas, desde enero he ido ganandolas limpiamente en la ruleta, los desafios, los cofres y los caramelos que transformas en regalos, ¡tenía más de 300 en total! Y esta mañana ya no hay ninguna, solo el martillo de piruleta que me ha dado hoy la ruleta. Al final borraré el juego y les diré a todos mis amigos que hagan lo mismo. Desanimais a cualquiera

**game - en - candycrushsaga - 5 - T16**

Addictive

I am so happy that you have so many new levels. I am on level 1,147. And when I liked there is still about another hundred levels or more for me to still get to. It was annoying when I had to wait for a new update when I would finish a set, but not anymore. Thank you for that, and keep the levels coming.

**game - en - candycrushsaga - 5 - T17**

Favourite time waster

Sweet calendar is flawed, only ever get to day two then it resets to beginning again... Love the sly look on bubble gum troll's face, very comical.. Game is getting better with more help in boost moves availability without needing to spend money... I think this game among many was created by Cain's people! Still best of the three candy games though.

**game - en - candycrushsaga - 4 - T18**

Thinking game

So candy crush is a great way to pass the time and get you thinking. It's a fun challenge to play especially when you're bored at work and shakes the old cobwebs off your brain. Some levels are harder than others the higher you get along the train track in the game but it's still a fun game. Can be frustrating at times but still fun nonetheless

**game - en - candycrushsaga - 2 - T19**

It has anxiety issues.

It doesn't let you think. It doesn't give you a chance to get out of sticky situations. I would have 2 moves left (and need only one to solve) and it would prompt be to purchase moves. What a waste and a joke. Spider Solitaire just sits back and watch you strategically win. Here you're presented with warped Candyland characters with sad faces when you don't cough up the green.

**game - en - candycrushsaga - 1 - T20**

Like the game but the app sucks!

Cheats you out of your bonus prizes and always disconnecting or take forever for it to connect. The prize calender just restarts its self whenever it wants to I haven't made it

pass day 4 in months.... sometimes the prize wheel doesn't even come up sometimes. And sometimes when it comes to the part to claim your prize suddenly I always lose connection and I miss out on the prize all together... like the game but very frustrating to play...

### **3. BOOKS**

#### **book - es - Cincuenta sombras de Grey - 5 - T21**

##### **Él MEJOR LIBRO**

La trilogía más exitante y conmovedora al mismo tiempo del mundo, te absorbe con la historia la forma de relatar de la autora, de verdad es estupenda recomendada 100% y esperando la película con ganas... el protagonista ideal es Matt Bomer el es el Christian Grey ideal.

#### **book - es - Cincuenta sombras de Grey - 4 - T22**

Me ha gustado bastante, pero hay que reconocer que la pobre Ana aún con toda su carrera no tiene mucha educación y además es desquiciante la mujer

#### **book - es - Cincuenta sombras de Grey - 2 - T23**

Malo

Es el peor de los tres libros. Realmente dejó mucho que desear y me aburrió tanto, que estuve a punto de no leer más de la trilogía.

#### **book - es - Cincuenta sombras de Grey - 1 - T24**

HORRIBLE!! No tiene nada: historia pesima, personajes estúpidos, mal redactado, ..ESTO ES ALTO VOLTAJE?? Si parece una película Disney. He visto más erotismo en los libros de Tom Clancy.

#### **book - es - Cincuenta sombras de Grey - 1 - T25**

Palabras que se repiten, la protagonista alucina con tocarle el dedo al otro, junto a un argumento forzado y personajes vacíos. Se lo recomiendo a amas de casa insatisfechas y niñas que creen que el amor son arcoíris y brillitos

#### **book - en - Fifty shades of grey - 5 - T26**

Addicting

When I read this book it felt like I was part of it. Feeling every emotion that went on. I cried during this book...I didn't want it to end I couldn't put it down read this book in one day then I found out that there were more books so as soon as I got done I moved on to the others I've read all of the book five times going on the sixth time.. Absolutely love them and advise everyone to read them.

#### **book - en - Fifty shades of grey - 4 - T27**

Good Story, Badly Written

I really enjoyed this book for it's story. The way it was written on the other hand can get really annoying at times, breaking me from the book. You can tell it's a erotic fan fiction...by the wording, etc...but I just couldn't put it down.

**book - en - Fifty shades of grey - 2 - T28**

Poorly researched, mediocre writing

I read first 150 pages in one sitting, even though the Grey character is super far fetched for his age. Then the contract part was just absurd... i put it down and never opened it again.

**book - en - Fifty shades of grey - 1 - T29**

GARBAGE!

I was wondering what everyone was so excited about. Took time to read. Big mistake! It sounds more like a 16 year old being brainwashed by some rich man. Oh and let's not forget the lovely message of "women will do anything for money" How much more offensive can this book get to its probably main readers. I could go on about how terrible it is, but I won't. Thank the all mighty it was a borrowed book and not bought. Your welcome

**book - en - Fifty shades of grey - 1 - T30**

Is there a word that goes beyond a waste of time?

An outing to the pictures was once fun. Excitement about the snack bar and the movie itself. Now mothers and fathers are taking their children to see this film. If you haven't given your children the talk. You're too late they have already learned too much. Your son is now going to be a rapist. While your daughter is going to end up buried in a shallow ditch all while you were too busy watching this smut not even realizing her killer is in the next aisle.

## **4. MUSIC**

**music - es - Adele25 - 5 - T31**

¡Me encanto!

Después de que se haya ido un tiempo y nos haya dejado sin nada de música, volvió y con todo. Todo este álbum es grandioso, tiene muy buenísimas canciones y eso es realmente genial, mis felicitaciones aunque Adele no vea esto:'v

**music - es - Adele25 - 5 - T32**

Vale la pena

No es necesario ser fan alocado de este artista para darse cuenta que Adele es música para los oídos y no solo ruido, al escuchar este álbum pensé, valió la pena esperar 4 años...

**music - es - Adele25 - 4 - T33**

Esta bien, pero no tanto como para que no hubiera sacado en tantos años nada. Esperaba más pero me gusta.

**music - es - Adele25 - 2 - T34**

mucho de lo mismo

**music - es - Adele25 - 1 - T35**

bonita voz mala composición

hello es una canción muy pesima

**music - en - Adele25 - 5 - T36**

This is an AWESOME album. I love the song Hello. It has a good beat. It is also a VERY popular song. I would recommend this album for people who love Adele. I also like the album 21 she cam out with, but this one is WAY better. I give this a 5 star because it helped me study when I needed to. I also listen to it when I just need some alone time. Adele is my favorite artist.

**music - en - Adele25 - 5 - T37**

Amazing

Adele is the best singer in the whole world and better yet I now have it on my phone and to all of those people out there right now that are saying that she is a let down... yeah you might think that it doesn't compare to the other albums she did...if you were a true fan you would love and adore her NO matter what kind of album she made. A fan is supposed to support the one they love and care about NO MATTER WHAT

**music - en - Adele25 - 4 - T38**

Soulful Pleasure

I love her music!!!! I can listen to her all day long. Her music is soulful..honest...clean. Most artist today music is full of GMO"S and FILLERS. NOT worth the time. My favorite jam is "Water Under the Bridge" Awesomeness Adele Rocks!!!

**music - en - Adele25 - 2 - 39**

To depressing, i do like the first track thou.. sorry adele, i was really excited about the new album then disappointed ):

**music - en - Adele25 - 1 - 40**

Overrated and overhyped.

This adele's music is so dull,boring,and depressing in fact it should have a warning advice sticker on the cd saying do not Listen if you suffer with mental health problems.this woman's voice is so very overrated.

## 5. FILMS

### **film - es - Frozen - El Reino del Hielo - 5 - T41**

Me encanta

BEs la mejor peli pero yo diria q no las vendieran por play estore xq hay niñas(o) q quieren verlas ino pueden ir al cine i solo lo ven por el mobil pero resulta q las vende llo les digo q mejoren eso graciasi tanvien lo digo por todas las peliculas i si algien mas me apollan gracias

### **film - es - Frozen - El Reino del Hielo - 4 - T42**

Muy buena

Me ha gustado mucho,porque es una película muy divertida,además te enseña el amor fraternal

### **film - es - Frozen - El Reino del Hielo - 2 - T43**

Flojisima historia y sobrevalorada que no vale la fama que tiene es aburrida y muy lenta y lo unico bueno es la animacion pero no justifica el bombo que le dieron

### **film - es - Frozen - El Reino del Hielo - 1 - T44**

Cacota

Frozen es una pelicula terrible. Transmite el mismo mensaje que todas las pelis disney: encuentra maromo y pa ti pa siempre. Hiper mega sobrevalorada. Y mucha Elsa pero casi no sale. Y ademas es una sosa amargada. Solo mola el pelo. Lo mejor es el muñeco se nieve. De verdad sois todos muy pesados con esta pelicula q no vale mas que para pasar una tarde y olvidarla despues. Por momentos se hace superlenta. Opinion objetiva de verdad. Trust me. A ver si teneis personalidad.

### **film - es - Frozen - El Reino del Hielo - 1 - T45**

No la aguantó

Es una película que lo único bueno que tiene es la música

### **film - en - Frozen 2013 - 5 - T46**

Amazing! Lovely movie.

Frozen is by far my favorite movie. It is funny,kid friendly, and a movie worth watching more than once. It has amazing songs. The end it to sweet and lovable. I loved how at the end it was not a guy, or prince, or something like that. It was Elsa true love and bond that saved Anna, not some guy, and that is awesome. Through this whole movie I was laughing my head off and in awe. The snowman, Aloph is to funny! He just cracked me up! I recommend that this movie is worth 14.99. I love this movie.

### **film - en - Frozen 2013 - 4 - T47**

JKiwi

I thought that this movie was a very well put together movie that is suitable for all ages. I think first, I should take the time to tell you that the animations are AMAZING. It is so clear and so bright that I will remain stunned at this work of art for years to come. It also must be watched in HD to get the full experience. The story is new.....ish. It has new characters and such but it also has the love story and plot structure that Disney is very famous for. It is about as different to any other Disney movie as Cinderella is different from Snow White. Or Mulan is different from Pinocchio. This is very much the 21 century Disney Princess movie that we all have been waiting for. It will not disappoint you. It reaches all of its goals, but it isn't revolutionary. If you have children you really have to let them watch it, if not, it is up to you.

**film - en - Frozen 2013 - 2 - T48**

Really really bland.

Story had its moments but becomes very bland and pointless. Follows many children movies and doesn't bring anything new to the table. Seems to be good for children but adults won't find much in this movie, at least I didn't. Music was more annoying than anything else and made the film seem like it was a day long adventure, don't see how this won an Oscar other than there were no other movies that had music in them. Definitely over hyped Disney junk and fails to meet anything they have done in the past.

**film - en - Frozen 2013 - 1 - T49**

Overrated

Overrated for no reason. Without let it go this would've been an ordinary movie. I saw it twice and was like what the hell is so fascinating about it. Anna is a moron, Olaf is a retard who thinks he's funny, the villain is predictable and Elsa is a Victoria Secret model with a Grammy award winning voice who owned the movie. It's not like Part of Your World or Reflection or Colors of the Wind that sang with uniqueness originality and emotion. The Lion King will always be 1 and Disney knows it.

**film - en - Frozen 2013 - 1 - T50**

WORST MOVIE EVER!!!

HATED IT!!! This movie was terrible about 2 minutes into the movie I wanted to destroy my TV tear my ears off and gouge out my eyeballs. I hated the music it stunk like butt. Love is an open door is repetitive. Let it go is annoying. Do you wanna build a snowman is pointless more like do you want to KILL a snow man. For the first time in forever is so freakin STUPID with a capital STUPID, and the song about reindeers and people was just plain weird. I would rate it 0 stars if I could. □□□□





## Appendix F. LARGER CORPUS

**Language:** en (English)  
es (Spanish)

**Application's name:** Clean Master (com.cleanmaster.mguard)  
Instagram com.instagram.android)  
Candy Crush (com.king.candycrushsaga)  
Angry Birds (com.rovio.angrybirds),  
All the Light We Cannot See  
Fifty shades of grey  
The Girl on the Train  
Avatar  
Frozen (Frozen 2013)  
Gravity  
The Wolf of Wall Street  
AC/DC (AC\_DC Back in Black)  
Adele (Adele25)

**Stars:** from 1 to 5.

**Review number:** from 1 to 200.

**Title example:**

<i>en-</i>	<i>com.cleanmaster.mguard-</i>	<i>1-</i>	<i>0</i>
language	application's name	stars	review number

### 1. ENGLISH TEXTS

#### 1.1. APPLICATIONS

##### en-com.cleanmaster.mguard-1-1

This app act like a virus.

Spam City! Even with notification off! Very aggressive with commercial. Reply: I turned off notification in the app of my phone ( in the setup ) And still got spamed by notifications...how you ask? When I plug it in the charger! Yes! More Spam from this app pop's again! Ok now they found a way to bypass the setup of my phone..wow...Since the update, I have to disable the battery spammer to ""off"" in the app. I know you guys need revenues but common, slow down with the spam.

##### en-com.cleanmaster.mguard-1-2

I liked this app but it has become invasive and has some issue

I liked this app a few years back, but it has since become littered with adds and the functionality of the app itself is comprised. Even after ""cleaning"" my junk, it will still repopulate the same amount of junk needing cleaning on a secondary scan. I am afraid to even use this app anymore, because it keeps deleting my contacts. I make sure to uncheck

anything associated with contact data, but it has happened so many times now and i cant afford to lose these numbers anymore.

**en-com.cleanmaster.mguard-1-3**

I don't think it does much.

Thermostat is always wrong. It said the CPU was 136°, but their other product, Battery Doctor, said it was 97°. Since I'm not burned, I can assume it's wrong. Can't turn off news in my notifications without turning on the notification cleaner I don't want running. It is determined to keep running something even when I'm not using it. I haven't seen much benefit but I've wasted a lot of time trying to get it working. Downgraded from 2 stars, I'm considering removing it.

**en-com.cleanmaster.mguard-2-4**

Used to like this app! Now its ruined!

This app is starting to piss me off! It runs slow now when cleaning junk files from phone. This app is meant to be a RAM booster anti virus and app lock but now its filled with junk features i dont care for like the dumb newsfeed and weather. If i want news or weather i can get that from a news or weather app. Thirdly the applock feature that used to protect your phone is no longer working and im now being asked to download another app- CM app lock. Too much junk and extra apps to make this app work.

**en-com.cleanmaster.mguard-2-5**

Constant notifications and false alarms

As a temp cleaner this app is still my no 1 choice as I haven't found anything better, but I'm getting sick of too many notifications and false alarms: by cleaning junk files or killing background apps the phone stops but the overheating notification does not go away, wasting battery. I also dislike it when I get pushed additional programs I don't want: I got iSwipe installed to my phone by stealth and I absolutely HATE IT BUT CANNOT UNINSTALL as it doesn't show in the list of apps!

**en-com.cleanmaster.mguard-2-6**

Battery drain I feel since adding theses apps my battery dies 3 times as fast. With the apps on battery save mode and with out me really using it.(a vedio during lunch, tunes for the ride to work and home) my phone would be at 50-60%. Used to be 80-85%. Actually useing my phone kills my battery in a few hours(music, Web surfing, one or two veido). and it keeps trying to get me to get more apps. Feels a little scamy.

**en-com.cleanmaster.mguard-4-7**

Was just looking to clean up my b phone

I was just looking to clean junk files off my phone and this app does it well. Searches all over my files and apps for things that are useless and just taking up space and slowing my phone down.. update. It's still a good app but it locks new apps from having permission to

things they need to work properly cause it thinks it's a virus or fraud. So to get things to work right you have to uninstall it then put it back later. Other than that it works

**en-com.cleanmaster.mguard-4-8**

Good app but data usage issue

App works well, only issue is I went to check my data usage and saw that this app had used up over 400mb in one day, solely in background data. What is causing this and how do I get it to stop?

**en-com.cleanmaster.mguard-4-9**

Love it... But

I really like this app, it is the best task manager I've had, my phone runs better and I am really maximizing my (limited) storage space. I just wish there would be an ad free version.

**en-com.cleanmaster.mguard-5-10**

I like the Clean master very much the best..

Thank you so much for the option I ask in the settings..... BACK TO 5 STAR !!!! Because this is the best cleaner for your HP..... thanks again people.... OOPS, the button for set off the news, we have now, like I say, formidable, BUT NOW PLEASE A WORKING BUTTON,... hahaha, this is really not working..... sorry.....

**en-com.cleanmaster.mguard-5-11**

It did work

I watched the you tube video first. My samsung 8 gig galaxy tab kept running out of storage, with a 32 gig inserted card, after a initial run then a restart and run it again my total free space of my 8 gig on the tablet was reduced by 2.1 gig pretty good results. It does seam to have the speed it had when new now also. I do notice a few adds but the device seams to be working great. It doesn't seam to have hurt anything. A good app I believe.

**en-com.instagram.android-1-12**

It's always giving me an error editing photos, and has no built in emoticons.

I don't have any emoticons because I use android and SwiftKey. Why aren't they just built in? My main complaint is it always gives me an error editing photos, and won't let me show the real size. It forces me to crop. But frequently I get errors trying to just post a photo. I'm guessing it doesn't like any type of panoramic shot (even a very short, barely-panoramic panoramic photo). Really frustrating. I can't even skip editing just to post, so there have been several photos I couldn't upload.

**en-com.instagram.android-1-13**

PROBLEMS: wish I could give this negative stars

I have to renew the link to Facebook everytime I switch accounts. 2. Ig automatically signs me out of my accounts for some reason saying. ""owner may have changed passwords"". 3. IG removes images without giving any specific reason that may have violated community guidelines. 4. Community guidelines include removing artistic images or drawings that show the ""demonic"" nipples but has no problem with bullying or racist images. 5. Ig terminates accounts based on the mentioned guidelines without any warning.

#### **en-com.instagram.android-2-14**

Miss the old Instagram

Really hate the nonchronological timeline. I mean, what is the purpose of it being nonchronological???? I feel that the instagram now have got too many features that I don't use and is a waste of space. Also, why must the location tag be decided only via GPS? I only upload my photos at home so it makes it very difficult for me to tag the photo's location....

#### **en-com.instagram.android-2-15**

Update Problems

The Stories update doesn't bloody work on my account, yet I wasted time downloading such a pathetic update which doesn't even completely work. The circle icon at the top of the status frame does not appear on my account, and it does on my mates. Prime example of a inconsistent update, algorithm malfunction. Besides that it's a good app. Just needs a bit of fixing here and there.

#### **en-com.instagram.android-4-16**

It's awesome but...

It's awesome, but for a while I didn't have a phone number but I had instagram after a while the thing popped up when you had to verify your phone number, my friend said that his just went away once he exited out but mine I didn't let me exit out, If I didn't have a phone number and verify I could follow people, post, like or comment. Please fix that for future people.

#### **en-com.instagram.android-4-17**

Its great!!!

It helps me keep in touch with friends and family that have moved away and is a way of sharing great photos and memory's with others. I ?? the new stories idea its great the only thing u would say is what if you saved all stories and not just jeep them in for a day its only a suggestion other wise excellent.

#### **en-com.instagram.android-5-18**

Awesome

This thing is this app is like so much but awesome I can't say anything about it but it's awesome you really want to buy it I will get it I mean I said goodbye. So I really think

that you should get this but if you really think that your kids shouldn't have it in less than 10 or older. But this app is really awesome so you should have it back if y'all are thinking about getting it I just wanted to let you know so you can have all the stuff information about it y'all can post pictures Anytime by the way it d

**en-com.instagram.android-5-19**

Love it

Instagrams fun, u get 2 watch other people videos and like them and people can do it 2 ur 2 also u can also talk 2 people on instagram u can have a private account or a open account if ur account is private only people u approve can see it so people u follow can see it and if ur account is open everyone can see it and if someone is bothering u on instagram u can report them or u can block them so on a download instagram and start makeing new friends and enjoy the fun!???????

**en-com.instagram.android-5-20**

Awesome app for sharing and viewing pics from around the world!

I love this app! It's helped me bring some passions and hobbies into my life I didn't even know I cared for. Sharing photography, food and travel has been fantastic. I also love the editing tools. This app is so much fun. Go out, live, take lots of pics and share them with the world!

**1.2. GAMES**

**en-com.king.candycrushsaga-1-21**

The sourest, bitterest, worst game around!!

Remove the wheel spin. Can NEVER win a jackpot. WHAT the hell do you have to lose if your players win the jackpot? Remove the ANNOYING owl. Lost a 35 moves level on the FIRST swipe. Remove the CRIMINAL 5 lives limit. Above all, SO MANY BUGS. 50 gold blocks went to 0 on doing an update. And now I've lost ALL the powers I had won. NO ONE even replies to the feedback. Such a SHAME & DISGRACE for a company like this. I'm never recommending this game to anyone. I am now at peace.

**en-com.king.candycrushsaga-1-22**

Very disappointed

This game is having 2 many problems. I had a striped booster pressed it so i could use it & the game acted like it was going to play & came back twice for me to press play again. The game took my striped booster without me playing it. Fix the bugs in this game!!! Also as of 7/24/16 this game is doing the same thing it was before when u play 1 world & have no more lives left then u have none in either world. Whats the point of having 2 worlds & 5 lives in each if u cant use all from each world. Very MAD!!!!

**en-com.king.candycrushsaga-1-23**

HUGE BUG!!!!

I love candy crush so much, trust me I wouldn't sleep until my life is finished but nowadays, ever since I updated the app, theres been quite a lot of issues. 1st of all it makes my phone so slow, when I turn it on it takes away my life without me even playing the game, it even restart my phone and like i said I reduces my life without playing the game. I don't understand why. All am just wishing is that I should never had updated the app. So please try and fix these issues then I'll rate it a 5 star

**en-com.king.candycrushsaga-2-24**

Five lives for both worlds???WTH And daily rewards disappeared

I love candy crush but after the new update why is it that my lives from the normal candy crush is tied to the ones on the dreamworld? I thought those were two different worlds and i can play dreamworld after my lives finish in the other one.. now i cannot. Its just 5 lives for both worlds.. that sucks big time. I have to choose which side i want to play at a time???? Also i play everyday and i had started getting the daily rewards then one dayni find myself back to the first without nothing.. c'mon guys..

**en-com.king.candycrushsaga-2-25**

Constant Glitch/Problem

I loved this game but disappointed due to a consistent issue. I love the special levels/events like Cake Climb or Treasure Chests and they are a big reason I play. Well, even though I get notifications saying they're available and those around me have them, nothing appears in my game. Twice now I have uninstalled and reinstalled or cleared my data and bam! the special event levels appear. The big problem with this is I lose ALL of my items when I start again. Edit: Reinstalled so we'll see...

**en-com.king.candycrushsaga-4-26**

Addictive, but...

The only reason I give the game 4 stars out of 5 is the poorly thought-through Daily Booster Wheel. While you can only spin the wheel once a day per account, the booster you win is not accessible on all devices. In other words, I can spin it and get a prize on my phone, but when I play from my tablet, I can no longer spin the wheel, and I do not the have booster I won on the phone. That's my only complaint. Otherwise, it's a great color puzzle game.

**en-com.king.candycrushsaga-4-27**

I love this game. Lost one star because waitin 3days to continue is very annoying. I'm not on Facebook 2 get help to nxt episode. pls try change that because not every1 is on Facebook. I'm gonna b waitin now 3days till I can continue as per usual. What is going on as I'm back to level 1 AGAIN and I was way over level 500. So totally annoying. Can't believe I'm back to level 1 AGAIN.

**en-com.king.candycrushsaga-4-28**

I still love this game. I only gave it 4 stars because i think you should have more than 5 lives at a time and i do not think you should have to have wait over 70 hours between episodes if you do not have 3 friends to unlock you or the gold bars to buy the next episode. No one wants to wait 3 days especially when they are having so much fun.

**en-com.king.candycrushsaga-5-29**

Awesome.e

It's such a fun game to play when I'm alone and no one will talk to me. I wish I had a friend. One time I cut my foot with a can. I also had a dog once. I like to hangout with the ants outside and follow there trails. I ate dog poop one time that had a ring in it. Fun time. But anyway this game is alright. I hope I can be cool like these people. I wanted to be a ninja but I could get into the school for ninjas in ninja turtle town. HOW you are not the only thing that has I don't know how you are a couple ki

**en-com.king.candycrushsaga-5-30**

GREAT / ADDICTIVE , BUT..

I love this game but not enough lives. ! I feel you should not have to ask friends for more lives. Maybe it could be where we earn them by getting coins / points by getting to next level & we use the coins to buy lives / power ups instead of asking for the lives or buying them with real money..! That would give the game another reason to win is to get the coins / points to upgrade. I like the improvements . .

**en-com.rovio.angrybirds-1-31**

Greedy Rovio. Way way way too many advertisements!!! Greedy Rovio, King of Ads!

At one time Angry Birds was a beautiful masterpiece with very few advertisements but then Rovio got increasingly greedy. Now the game is borderline unplayable. There is a video ad every time you retry, every time you change level and if that wasn't enough, during the game, video banner ads appear at the top whilst you're lining up a shot. AND, annoyingly, video ads play when not connected to WiFi which means as well as Greedy Rovio making money, you're essentially paying for it. -5/5

**en-com.rovio.angrybirds-1-32**

Powerups

There are levels you need power ups on to get there stars. That is really dumb. This game is totally overrated. Also I wonder how many ratings are made by Rovio. ActuallyI've gone from hating this game to actual astonishment at how they can do anything that is obviously for nothing but money and it still has any stars whatsoever. This is a sickeningly obvious way to force you to spend money that still lies unseen by the pigs that are the ones that rate this game higher than one star.

**en-com.rovio.angrybirds-2-33**

Love the game, hate the noise

I understand the price I pay for a free game is ads. I don't quite understand why Angry Birds runs a full-page ad immediately following a video ad, but I can accept it. But it is rude, indifferent, disconcerting, annoying, entrapping to run a video ad FULL VOLUME PLUS. The mute button usually doesn't respond so what can be politely described as awful music assaults my ears. Worse, though these ads usually run every three ads or so, sometimes it's as often as often as every play.

**en-com.rovio.angrybirds-2-34**

Ads ruin it

The game itself is a lot of fun, and the micro transactions aren't too in your face or necessary so you can consistently win without buying them. The reason I'm giving two stars is because the ads are horrid. They're far too frequent, most are video ads so they needlessly eat up large amounts of data where a simpler ad would have sufficed, and the ads will randomly revert to max volume despite me repeatedly turning the game to silent. There are better games without such huge drawbacks, free or not.

**en-com.rovio.angrybirds-2-35**

Fun game but broken with restricted profiles

My daughter loves this game but it doesn't work properly with restricted profiles. If you purchase the ads-free option in the main profile (\$1.99 in-app), that purchase doesn't carry over to restricted profiles. So my daughter is still subjected to the ads and I wasted \$1.99.

**en-com.rovio.angrybirds-4-36**

IF...

Possibly the most awesome game ever. Very clever, addictive, and entertaining. EXCEPT... Those stupid #@\$ banner ads pop up right when you are going to or right after launch so birds can't be seen or manipulated. Multiple 15 & 30 second ads after EVERY play. Pathetic. Place those #@\$ banner ads after game play where they don't interfere & cut down on the big ads. Banners are enough. Excellent game and love it but right now it's an irritating make you want to just Bowl instead game. Totally sucks. FIX IT.

**en-com.rovio.angrybirds-4-37**

Awesome game

I like this game because it is soo fun to knock over the piggie fort. Once again its too cool to rate. If it was unlimited rating to 100, I would rate it 99.5 Well done Angry birds Creaters. You made a very well liked game and I know it will keep growing because this game is awesome. P.s I have installed all the other angry bird games and I love then. If you are reading this then you should totally check out all the angry bird apps

**en-com.rovio.angrybirds-5-38**

Fantastic and original, a superb game. (Recommend)



They don't call this the no.1 app for nothing, it's the original and best clean game about. One needs to work out how best to strike the piggy structures by flinging the birds in a sling (funny though, I thought they had wings). It's a novel idea for a game and will no doubt continue to appeal to a wide audience for some time to come! (It's as big as what the old lemmings game was in the 1990s) It's so addictive and fun to play whilst having a good feel about it. For me beats Angry Birds 2 game very easily and will no doubt remain in the future. However, the new eagle challenge is a crude way to fork more money from you on which prices for the tickets are beyond belief! (You have been warned).

**en-com.rovio.angrybirds-5-39**

Cool game

Hi Rovio, I love this game (I thought the movie was excellent). I think that although there are a fair few adverts that doesn't matter because you can skip them after like three seconds. I'd understand why people are annoyed if the ads were at least a minute long and you couldn't skip them but they aren't anything like that so people stop complaining. The idea is really imaginative and I like how there is a story to it all.

**en-com.rovio.angrybirds-5-40**

One of the greats!

This game is in the same league of great game franchises like Super Mario, Mega Man, Legend of Zelda, Sonic the Hedgehog and more for a good reason. This game is addicting, lots of replay value, and still stands the test of time. This game is a must for gamers. Even if I have this on my Android, I still will purchase Angry Birds Trilogy for the Wii as well.

### **1.3. BOOKS**

**en-All the Light We Cannot See-2-41**

Although the characters are strong the book was a bit dry but well written. Personally I prefer to read to get me away from the harsh realities of life and WWI

**en-All the Light We Cannot See-4-42**

I don't know if this is actually a five star book or if I was just disappointed because of all of the hype. It's a good book. The story is actually very good. The writing is great. That might have been the problem. It was a little too great. It distracted from the story. Reading All the Light we cannot See is a lot like reading poetry. It makes you think. But there were times when I lost track of what Anthony Doerr was saying. It was a little over the top. But I wouldn't let that discourage you from reading the book. By all means read it!. You'll enjoy it. I personally could have done with less of the fancy footwork and a little more story. But for all the poetry lovers out there, this is your book.

**en-All the Light We Cannot See-4-43**

Get ready to be surprised

Wow, what a great read, not at all what I was expecting the book to be. I was prepared for more easy reading fiction, but the realism and the fact that the author portrayed life as it really is (happy, sad, easy, hard) was refreshing. This book is dark, and real, an adult oriented novel. Really good.

**en-All the Light We Cannot See-4-44**

Marvelous

Compelling story with beautiful characters and something descriptions. It loses its momentum toward the end -- something I don't fault the author for necessarily because hope do you effectively come down from the tragedies of World War II and the brutality of the Nazis? A most worthy read.

**en-All the Light We Cannot See-5-45**

Such a beautiful book

Doerr writes gorgeously. That's the first thing. Once you get past that, you'll find that this book have wonderfully sympathetic characters and an intricate story line that is a mix of history and a bit of fantasy. I found the ending a bit abrupt, but that could just be because I miss the atmosphere of Doerr's war tale. This has become one of my favorite books and I definitely recommend it to all.

**en-All the Light We Cannot See-5-46**

Amazing

This story is remarkably told, in a poetic way attached with metaphors, it almost makes you feel what's going on in the setting instead of just reading. I always think the title was also chose for this book because of the way we can ""see"" the setting in this book, but we can't physically see it. Truly is a work of art. This is one of the books that I will probably read more than once!

**en-All the Light We Cannot See-5-47**

Loved it! A book I am proud to say sits on my bookshelf (when it is not being borrowed of course).

A beautifully written story. It took me entirely into the somewhat atypical worlds of the two intriguing and wonderfully constructed main characters. In summary, a great read from beginning to end. Definitely a book that I will buy again for gift-giving. :)"

**en-All the Light We Cannot See-5-48**

Review: All the Light We Cannot See

Just wow. Probably my favorite, and perhaps even the best, book I've read this decade. I went in with high expectations and they were easily surpassed. Intricate, balanced, brilliant story told with beautiful prose and powerful imagery. Just read it. Seriously.

**en-Fifty shades of grey-1-49**

Poorly written

The story is absurd and disturbing. For the author to try to make the reader sympathize and love a man who manipulates a virgin into s&m is scary. People are into that, yes. But the most experienced will tell you that's something you explore with someone you know well and trust. Not to mention there's a borderline rape scene. I mean some parts honestly read like a rape apologist novel. It's full of stereotypes that men use against women. If you want erotica, even kinky, there are some good novels out there. This is not it. And I'm sad to see so many teens reading it thinking this is what sex is about.

#### **en-Fifty shades of grey-1-50**

Love the title.....

Read the first two pages to see what was so great. I was bored. It seemed familiar, though. I'm glad I didn't read anymore. Buying her place of business so he can keep an eye on her. Putting a GPS tracker on her phone right after they meet. The mental and sometimes physical abuse. How can this be romantic. He's a psycho. Is this the type of man you want you're daughter involved with? And, yes, I didn't read it all, but there are plenty of quotes all over the Internet. Have more respect for yourself.

#### **en-Fifty shades of grey-2-51**

Really, Bestseller?? (My Opinion)

I am super curious as to how this book could ever be a bestseller. Clearly simple minded people read this because not only does it have no excitement other than the horribly written sex moments, it is way too repetitive, and has a very small vocabulary. People, there are ways to make sex actually enjoyable to read. This is not one. It is not interesting, it is not well written, it is not exciting. It is repetitious, typical, poorly elaborated, and lacks climax. Conclusion: Author needs work.

#### **en-Fifty shades of grey-2-52**

My review...

For those who wish to read it: I didn't think the books were poorly written in and of themselves, but I did not care for the fact that as the books progressed, Christian Gray became more and more emasculated to the point I couldn't stomach the third book. I'm thinking to myself, why?! Anastasia is so plain Jane and ugly that I question the guy's taste in women. I had zero interest in Ana as a character, and I wondered when Christian was going to wake up and find someone with an IQ above 12.

#### **en-Fifty shades of grey-4-53**

Enjoying but...

It's unrealistic that this ""virgin"" girl is so skilled in her first sexual encounter and why does the author consistently have the characters 'murmur' and 'mumble? ' I'm having trouble entering the story with the misuse of the words. Also, why does she initiate meeting Jose' for 'coffee' when she doesn't like coffee? I'm early in the reading, so maybe I'll change my rating if I find a reason that outweighs the aforementioned issues. Ok. I'm intrigued. Next book!

**en-Fifty shades of grey-5-54**

AMAZING

At first, I was a little skeptical. I read this book because I was curious to see what the commotion was all about. When I started reading the book, I was a little freaked out both good and bad ways I guess. But as I kept reading, it was so intoxicating and just addicting. I LOVE THIS BOOK. I couldn't put it down and every thing I did or as I went on with my day, I just kept thinking of this book. It's so good. Grey is smoking ugh. I'm reading the first book in Christians perspective rn and I love it.

**en-The Girl on the Train-1-55**

Depressing

I found the protagonist' line of thought astonishingly irritating. She is weak, insecure, an alcoholic and depressed. Much of the plot revolves around how the protagonist thinks how pathetic her life is. The quality of writing is not astonishingly good either. So I hated the plot, hated the character, found neither inspiration nor insight in the book and did not particularly enjoy Paula's writing. Maybe I would have liked it better if I could relate to the character more.

**en-The Girl on the Train-1-56**

Boring

No idea what the hype's all about. The plot was derivative and the characters dull and thoroughly unlikeable. That the main character keeps making the same mistakes and can't just leave things alone is really irritating. That said if you want to read a story about an alcoholic loser who nearly gets herself killed and hasn't grown as a person in any way by the end of the novel, jump right in. Seriously awful.

**en-The Girl on the Train-1-57**

The Girl on the Train

Very disappointed. How this became a best seller I shall never know. The ""Girl"" seemed to be permanently drunk & it was well into the book before the crime took place by which time I was just flicking through pages & gave up."

**en-The Girl on the Train-2-58**

I have the sample of this No 1 best seller. I'm into this crime, mystery, thrillers fiction novels big but. We all know how a movie starts slow & seems to drag on forever before getting better well

This book feels as if it's the same way. I can't get past the first or 2nd chapter it drags on.

**en-The Girl on the Train-2-59**

Ok

I didn't understand where Hawkins was going with the culprit in places and I didn't see the point of certain characters and events. Not really any suspense and the storyline was

very predictable. I did like the diary style and it was good to read the perspective of the different characters.

#### **en-The Girl on the Train-4-60**

Simple language, complicated story

Really easy to read due to the simple language, however the twisted story made the casual people lives exciting. A good psychological explanation behind each character's actions, which makes you feel sympathy for everyone. Consequently, it's impossible to identify a villain or a hero as everyone had their own reasonable motives.

### **1.4. FILMS**

#### **en-Avatar-1-61**

Really hated this movie

I really hated Avatar, the special effects were incredible, but the storyline was crap. I was in the military during the 60's and a lot people hated us, calling us really hurtful names like, baby-killers and this movie stereotyped the military as murders, mindless goons, thugs who just go out to kill. I'm pretty certain that Avatar 2, 3 and 4 will portray the military the same way, even though military personnel are dying for us overseas.

#### **en-Avatar-1-62**

More Crap Cameron

Cameron is the new S. King. Overrated, arrogant, proselytizing sob story SHITE. Avoid this Crap at all costs. P.S. Eric de Dongqueer needs to remove his head from Cameron's rectum.

#### **en-Avatar-1-63**

Boring

I have an idea for special effects shots, now let's write a story to go around them. Animation looked incredible, even more so considering they used lightwave, but the movie just plain sucked.

#### **en-Avatar-1-64**

So overrated

Storyline unoriginal, flat and down right boring. Unobtainium? Really? Poorly written, poorly produced, poorly acted. Just downright bad Science Fiction. Mediocre antimilitary propaganda film if that's what your looking for.

#### **en-Avatar-2-65**

Hands down, the most overrated film I've ever seen.

I don't understand the hype. The cgi characters are all ""rubbery,"" like PlayStation 2 (maybe 3) cutscenes. And, the 3D (I saw it in IMAX) was most definitely NOT

impressive. The story is a sci-fi ""Dances with Wolves"" which would be more impressive if it didn't feel so stale. Nothing new here.

#### **en-Avatar-2-66**

Without paying attention to the graphics...

... three words sum up the entire breadth of the plot: Pocahontas In Space. The only thing going for this movie is the special effects, period. It won't make you think, it won't question your beliefs, and it barely deserves to be considered science fiction.

#### **en-Avatar-2-67**

Good special effects, otherwise terrible

Visually, it's impressive. But the characters are flat and unlikable, the main protagonist is the guy who should have died in the first hour, and there's no subtlety when it comes to the environmental message or who's going to end up a villain.

#### **en-Avatar-4-68**

Fun and Epic

A likeable romp in a mystical alien world. Watch it for the emotion and the beautiful world that it's set in because the story is pretty cliché, hence my 4 star rating. I'd have given it 5 if it didn't feel like Pocahontas meets Alien VS Predator.

#### **en-Avatar-5-69**

Avatar is one of my all time favorite movies

This movie really show what the European people have been doing all over the earth to people of color and people of the land. The european have destroyed wood land, forest, and whole communities for his own greed and pleasure. Yes I know some of you may say what I'm saying is racist but history will show that it's not that what I'm saying is fact! This movie is great. It also show how the european and his armys go into a place that is not theirs and take what does not belong to them and inslave the people also labeling them as savages not worthy to live according to way of live. Take a deeper look and go back in history and you will see it, you will see how the Europeans use their technology and machine to kill peace loving people, how they come in and try to enforce their ways on the peace loving people and if the peace loving people don't bow down then the european call them names and then set out to kill the land and people living on it. This movie also show how the people are connect to the land and all living things around them, how they and nature are one I also think if the writers would've made the avatar African, Native or any indigenous being, people would've walked ou

#### **en-Avatar-5-70**

Fun and heartwarming and colorful It shows the passion between the alien girl and the ex-marine. And even if he was torn between two worlds, jake sully helped the people and with that he betrayed his kind. But besides that this movie is amazing, and I loved every minute of it. The special affects were my 2 favorite things about the movie because the

affects show what kind of environment the people live in, they show what the world looks like, and how amazing at how much color there was. WOW!!!

**en-Avatar-5-71**

That is the most fantastic movie I have ever seen in my life. It is the tenth time I am watching since it came out in 2009 and I can't get enough! The world that James Cameron has created is so stunningly beautiful! Every time I watch it, I keep noticing new details in graphics and in the story as well. Some people say that it is simply. It definitely isn't. You just be careful while watching it. It is a pure masterpiece. I recommend it wholeheartedly!.

**en-Avatar-5-72**

"Just all around amazingly beautiful movie Best movie on the planet for many years to come everyone in the movie industry knows it can not be beat no matter what haters say. James Cameron and everyone in part with Avatar have done beyond fantastic in every way on this movie . I can not wait for the other 3 movies to come. Love your work everyone in Avatar best wishes to health wealth and knowledge now plz lets all start recycling ;). Lots of love"

**en-Frozen 2013-2-73**

It's okay.

This movie has its goods and its bads. When I first saw the trailer, I really wanted to see it and my brother got the movie through red box. I was really happy when he got it, but after seeing it I was pretty upset because I thought the movie was as great like everyone had been saying. It is not the best since the Lion King but I did like the family is true love twist. And I don't know why Frozen fans get hot & bothered when people don't like the movie, Disney gets paid either way & y'all aren't getting paid.

**en-Frozen 2013-4-74**

Cute And Catchy But...

The characters are cute, fun and the villain was a surprise. However, this is one of those films that is getting run into the ground for being so amazing. It IS an amazing film! But it lacks something... A lot of it did not make sense story-wise. I was perplexed as to why the so-called ""rulers"" of a land were NEVER involved in any political decisions at all. And it was not like they had a steward who did everything, they just didn't have anyone. Elsa and Anna talk about trade with other kingdoms but where did they learn to do that? Their parents have been dead for years and they have been literally locked inside a castle during that time. They only open up the castle for Elsa's coronation (why was she not crowned after her parents died? Who ruled in their stead?) and Elsa brings about perpetual winter in the wake of her anger at the naivete of her sister. But they were both locked in the castle for nearly their entire lives... Of COURSE Anna is going to latch onto the first guy she even slightly clicks with. Then we see an amazing musical sequence and

the movie progresses forward, for the most part, quite smoothly. Kids love this film so buy it if you have any. Rent otherwise."

### **en-Gravity-1-75**

So flawed!

It is hard to believe that the technical consultants were so asleep on this one: Surely, no one from NASA had anything to do with this. Right from the beginning the technical errors start to mount, and keep right on building to end: A mountain of bogus occurrences and, an impossible splash landing. The first big mistake is right at the opening scene. You see the earth rotating west to east but, then you see the space shuttle coming from the east and heading west, counter orbiting the earth's rotation. The earth is spinning at about 1,040 miles per hour at the equator; and along the equator is where the Hubble sits in it's orbit. No one is going to approach the Hubble space telescope from the east to match orbit, unless they already passed it up, with the waste of fuel in recovering it's orbital position and matching it's velocity. The last scene is just as bad. The reentry module should have burnt up during reentry; it isn't just going to fall into an ablative heat shield down position all by itself; it would have kept tumbling till it ruptured and disintegrated. Worst flaw? A poor script that top actors like Bullock and Clooney can't fix.

### **en-Gravity-2-76**

Most Overrated Movie Of All Time

Now sure, I get there's room for personal opinions... but come on, how is the entire world saying this movie is so perfect? I watched it and was thoroughly disappointed. From a technological standpoint, this movie is flawless. Regarding the story, at some point, the film makers forgot they were making a movie and were convinced they were making a space simulator. I'm sorry if my idea of fun isn't watching Sandra Bullock spin around uncontrollably in space for 20 minutes before anything else interesting happens. Her acclaimed acting literally consists of her screaming and crying like a scared little girl for half the movie, and then out of nowhere she becomes this strong, independent woman, because that's what audiences are obsessed with these days, regardless of inconsistency. There's very little in the way of plot twists or developments in this movie. There's a space accident, and she has to get back to Earth. Another astronaut tells her how, and she does exactly that. The end. The entire movie relies on the viewer's fascination with space to keep them interested. George Clooney acts decently well in this, but doesn't have a very exciting role to work with.

### **en-Gravity-4-77**

Great Movie, Average Acting

Yes, the physics aren't perfect. You want to know why? Because they filmed it on Earth!! It's not possible to get zero-G effects perfect on a planet that, does indeed, have gravity. They did the best they could and pulled off the physics, especially in scenes that were entirely CG. It was a great movie that had great balance and terrific special effects. The



acting however, was average. I am not a fan of either, and they didn't shine, but did good enough that it didn't detract from the film. Solid 8/10.

#### **en-Gravity-4-78**

Stunning effects and a taut thriller, but a fundamentally tired plot.

Gravity works excellently as a thriller, and the visual and audio effects are brilliant. Questionable science and laughably implausible scenarios are forgivable in service to the plot, which successfully strings the viewer along from crisis to crisis, but ultimately it's the setting that makes Gravity notable: the isolation and cold hostility of space. The plot at its core is one we know well: tortured by her past, the protagonist must confront her demons and find life is worth living.

#### **en-Gravity-5-79**

Wow!

It's been a long time since I've been so gripped by a sci-fi adventure. I found ""Gravity"" to be far more compelling than ""Interstellar,"" partly because the story is more tightly focused and partly because it's on a more human scale. The predicament of the stranded astronauts is genuinely terrifying, and their attempts to save themselves are moving. I was close to tears a couple of times. It took a while for both Clooney and Bullock's characters to settle down. At the beginning neither is believable as Nasa astronauts — Clooney is too much the fly-boy jock and Bullock is too much the nervous tyro — but as the film progresses both characters become more believable. The special effects are just amazing. Planet Earth is as much a character in this movie as any of the humans involved. I'm \*very\* glad I watched this, and I'd definitely watch it again."

#### **en-The Wolf of Wall Street-4-80**

Disappointing

While there was a lot of hype for this movie I found it uneven, overacted and way too long. The nearly three hour movie stars Leonardo DiCaprio who is much better in Django Unchained than he is in this movie. He is almost overacting in this role which is based on a real story and character who makes a cameo at the end. The sex, nudity and swearing goes way overboard. Believe me I love movies like this and Martin Scorsese is one of my favorites but this one is not.

### **1.5. MUSIC**

#### **en-AC\_DC Back in Black-1-81**

Back in black

The mini album I bought for \$3.99.doesnt play.when I try a window pops up tells me, cannot stream temporarily unavailable. What a waste of money.what a rip off.

#### **en-AC\_DC Back in Black-5-82**

What can I say???

I was going to write a review telling the whole story about how AC/DC are the greatest rock n roll band in the world, how they had the greatest frontman to ever take the stage Bon Scott. How he died in 1980 and rock music took it's worst blow since Hendrix. How the band so nearly decided they couldn't continue but carried on and they found a new equally unique vocalist, a geordie and his own man who gave it his all and together they made a landmark rock album. But I will just say this. Buy it,play it now.

**en-Adele25-1-83**

Keep your Problems to Yourself

Has Adele failed to learn the principal lesson that to burden others with your problems is impolite?? The next time I hear her complaining about one thing or other I just might get the inkling to buy her publisher out and close her career for good. What a colossal selfish pain in the tuckus. Leave music - NOW.

**en-Adele25-1-84**

Great personality, great voice, terrible music

More of the same, no progression since the first album. If you're a middle class white mum who's music sphere is narrow, this is amazing. If you have a brain you'd avoid it

**en-Adele25-1-85**

Too LOUD Adele

In my opinion, I feel that the loud almost shouting of songs, that seems popular with many artists at present, is unnecessary. Sorry Adele, but for that reason 25 is one I'll pass on this time.

**en-Adele25-1-86**

What ! 2 good songs only!!! What..

This album was all wrong should not have been out! Thee amount of songs that suck is just to many to count.. This is my openion. This should be free.

**en-Adele25-2-87**

"Love Adele, but I was disappointed in ""25"". Nineteen and Twenty-one were far superior to this release. It sounds very repetitive and boring after the first, most popular songs. One song could run into the next, and you cannot hardly tell the difference... Powerful, soulful, but monotonous. I feel that if this was Adele's first album, she would not have made it to the superstar she is today. I am still a big fan, but sorry I did not like this album."

**en-Adele25-2-88**

OVERRATED: Great voice. Subpar songs

Her voice is incredible. Her songs are not (with the exception of Hello & When we were young). Album's 19 & 21 were a lot better. I do not feel like replaying many of these songs.

**en-Adele25-2-89**

Disappointed

I LOVE Adele, have both of her previous albums and I still play them. This album I'm upset I bought, barely any songs that stick out and it's just like singing the same old song over and over :-/

**en-Adele25-2-90**

For die hard Adele fans only

Can't deny Adele's vocal quality and production superb..but it covers over the lack of really strong songs we have been used to from such a great artist. Shame after so long a wait.

**en-Adele25-2-91**

Amazing voice. Disappointing album

Was only a few songs that were so unique with the Adele edge. The few I want to replay. The rest were meh, forgettable. Disappointing. Of course she's still a talented singer.

**en-Adele25-4-92**

A good album with some great tracks

As someone who did not care much for Hello (guess I will be in the minority here), I was pleasantly surprised by some tracks - especially Million Years Ago. That one track alone justifies the price for the album

**en-Adele25-4-93**

Nice

The only reason I gave it a four and not a five is because I don't like when she sings fast songs. Lol I love her music and she sounds more beautiful in person.

**en-Adele25-4-94**

Not as good as 21

There are a great couple of tracks, but most are forgettable. Was expecting more from this album, it's not in the same league as 21 imo.

**en-Adele25-4-95**

Great sounds is back

Hello introduces this album powerful and strong...then it kinda drops off, still great but ""hello"" blows everything out of the water

**en-Adele25-4-96**

She has a beautiful voice, full of emotion. It's perfect music to have when your by yourself and need to reflect on life and love.

**en-Adele25-5-97**

Awesome voice

First of all, this is important, the similar artists that Google lists should NOT include Katy Perry or Taylor Swift. Those two could only dream of having a voice like Adele's. KP & TS cannot sing well live to save their lives. The only reason they sound OK on their studio albums is because of auto-tune. If you don't know what it is look it up. KP & TS could never pull off what Adele does singing live or what Lady Gaga did with the Sound of Music montage live. And those talentless bimbos are millionaires.

**en-Adele25-5-98**

NO ONE MAKES ME FEEL LIKE ADELE'S MUSIC

That's a pretty huge statement too, with all the musicians in the world, but it's TRUTH. I have to admit, it was HELLO & 2 close gf's who showed me HER light, and I'm completely taken away to a million places every time I listen to her voice, her songs. She makes me feel EXACTLY AS I DID those years ago. Others have done that too, but ADELE DOES IT BEST. THANK YOU ADELE, BLESS YOU AND BRAVO! A brand new fan!! Love you, Adele! ????????????

**en-Adele25-5-99**

Amazing!

I almost think that Adele had to pay the negative people on here to say something dumb just so we don't think there is a hex on us people for loving this women so much! Every album my appreciation for Adele grows, more and more, deeper and unwildly. I am truly thankful for artist who write songs that have meaning and expression. Don't get me wrong because I love some pop hits and the club bangers but every once in a while I want to feel like there is something more, a story, a really pure song! Amazing!

**en-Adele25-5-100**

Adele is awesome

I thoroughly enjoyed this album. I can not say enough about Adele's talent. Since her first album I've been captivated. Having been around the same age as each album, I feel as if I've grown with them. The young love heartbreak child of 19, the I am woman hear me roar for love and life 21....and now 25. It has a maturity, a soul, that I may have appreciated as a teenager but now older, I can understand . Buy the album it won't disappoint.

## **2. SPANISH TEXTS**

### **2.1. APPLICATIONS**

**es-com.cleanmaster.mguard-1-101**

Vaya que la arruinaron.

Antes era funcional y sencilla, ahora pusieron más funciones y a simple vista parece que ya no es limpiador, para que quiero el antivirus, las noticias, el iswipe, el applock, se supone que es limpiador, no app multiusos, la misma app pesa mucho, paso de pesar apenas 10 megas, a 68, los anuncios aún siguen siendo un problema, y de pilon al desinstalar juegos, me irrumpe con su ventanilla asquerosa, me puso, no se detectó basura, y encima te pone juega algo entretenido, recomendando juegos de matanza, QUE CREEN QUE ESO ES APTO PARA NIÑOS? PUES NO, He borrado muchos de mis juegos, sólo tengo 2, Ah, pues en el optimizador de juegos, gracias a esto, su estúpida app me pone: 1,2,3,4,5,6,7 anuncios, enteros, esa son 7 razones por las cuales debo desinstalar, pero si sigue, ahora si lo desinstalo.

#### **es-com.cleanmaster.mguard-1-102**

Cada vez peor

Antes funcionaba de lujo ahora tiene demasiadas cosas y a simple vista parece que limpie menos y sigue el fallo con la temperatura. Para que quiero las noticias, no es una app de limpieza, cada vez consume mas recursos esta app que otras, de apenas 20mb a estar ya ocupando 160 mb y encima ella misma no se limpia, empezó muy bien acaba mal, no tardare en quitarla. Ya he encontrado una que hace lo que dice limpiar y sólo limpia nada mas, y por cierto, elimina mas que la vuestra.

#### **es-com.cleanmaster.mguard-2-103**

Ilogica

Al principio de año la aplicación funcionaba perfectamente, ahora con tanto añadido absurdo y tanta actualización han perdido sentido lógico de lo que se supone es la aplicación, especialmente el boostcharge es un fallo total, solo colapsa el teléfono, vuelvan a lo simple, siempre es mejor

#### **es-com.cleanmaster.mguard-2-104**

Es una porquería de app no me sirve para nada

Se me pone muy lento el celular se me pega tengo q limpiar a cada minuto tube q eliminar casi todas las aplicaciones q tenia es una porquería. Esto es un regalo ahora lo boy a eliminar no sirve.

#### **es-com.cleanmaster.mguard-2-105**

Más publicidad? No, gracias!

Utilizando esta aplicación desde hace tiempo, ahora noto que ha crecido enormemente la publicidad invasiva. Me empiezo a plantear desinstalarla. Esas alertas que me saltan últimamente a pesar de tener la opción desactivada me están empezando a cansar.

#### **es-com.cleanmaster.mguard-4-106**

A cada rato molesta

Una notificación cada 10 o cuantos minutos aparece una notificación, que si esa app, que esto, que lo otro, que se yo, lo que tú quieras. Además, siempre se que quiere que instale

el antivirus (que es otra app aparte) y las publicidades siguen arrasando y eso le quita estrellas. Por otra parte, agradezco, ya que tenía 7 gigas usadas, en menos de 3 minutos me había desecho de 3 gigas enteras, en serio, no se que haría sin esta aplicación.

**es-com.cleanmaster.mguard-4-107**

Me gusta, pero hay un pequeño error

Cuando uso charge máster que ayuda a cargar más rápido y cerrar aplicaciones que gasten energía, me sale un bug, cuando salen las notificaciones cuando prendo el celular y lo deslizó vuelven a aparecer como 3 veces y no puedo desbloquearlo, e incluso en veces cuando deslizó estas ventanas de notificaciones de desbloquea sólo el celular, espero que porfavor reparen esto. Si arreglan el error les doy 5 estrellas.

**es-com.cleanmaster.mguard-4-108**

Práctica, conveniente y funcional

Estoy satisfecha, el app logra el trabajo. Ayuda a liberar espacio de memoria RAM fácilmente y enfría la CPU con rapidez. También ayuda a ahorrar batería con el modo ""Instahibernar"", que pone las aplicaciones activas en un plan de bajo consumo. Su interfaz es amigable y sencillo de usar. Es todo en uno... Muy conveniente. ¡La recomiendo!

**es-com.cleanmaster.mguard-5-109**

No y si descargen

Todos pensarán esta app funciona pero en vdd la app solo sirve con dispositivos lentos un ejemplo cuando una persona compra su dispositivo nuevo y descargan la app el dispositivo se lento automáticamente poco a poco en menos de 5 días lo que les trato de decir es que solo descarguen esta app cuando tu dispositivo en verdad es lento o medio lento SI NO es LENTO TU DISPOSITIVO NO DESCARGAR por que el se hará lento

**es-com.cleanmaster.mguard-5-110**

Súper con algún que otro detalle menor.

Súper útil la aplicación, lo único que molesta un poco son todas esas utilidades extras. Se entiende que quieren mejorar agregando cosas, pero la verdad no hace falta, engordan la app. En todo caso den la chance de activarlas no que uno se tenga que tomar el trabajo de desactivar. Simple sugerencias. Por lo demás 10 puntos.

**es-com.instagram.android-1-111**

La última versión de Instagram es una basura y no funciona bien.

La última actualización de instagram es una basura desde que actualice la app no me deja ver las fotos de las personas que estoy siguiendo me salen borrosas y no cargan, si quiero ver las fotos tengo que meterme en Google y verlas a través de Internet. Sólo pido que solucionen esto de una vez y que Instagram vuelva a funcionar como antes, como si vuelven a incorporar la antigua versión que no había problemas y funcionaba de maravilla, espero que solucionen esto lo antes posible, GRACIAS.

**es-com.instagram.android-1-112**

Si me gusta la app pero tengan cuidado

Tengan cuidado con un usuario que se llama alexispsylocibe es un usuario que anda buscando seguir a niñas o jovenes mujeres para mandarle imagenes de sus genitales y les queria advertir porque anda cambiando su nombre de usuario antes se llamaba alexisviajero25 y ahora asi se llama para q tengan cuidado con ese hombre y creo que se preguntaran como se que cambia su nombre de usuario a pues es que lo tengo bloqueado en cuentas bloqueadas asi que mucho cuidado para que no les valla a pasar lo mismo que a mi.

**es-com.instagram.android-1-113**

Valiente mierda de aplicacion

Esta aplicacion no vale nada, me da verguenza me a bastado con tenerla horas y ver que no sirve para nada yo creo que deberias de quitar esta aplicacion, porque a mi se me cae la cara de verguenza de ver a niñas algunas menores de edad enseñando sus cuerpos(culos y pechos..ect)creo que eso no deberia estar permitido, la mitad de esos padres no sabran lo que sus hijas suben y lo peor son los hombres comentando, algunos comentarios bonitos pero otros..mejor ni escribirlos. Deberiais tomar medidas y revisar fotos.

**es-com.instagram.android-2-114**

Me encantaba pero...

Ahora se copiaron de Snapchat! Yo tengo Snapchat (que los efectos son muchos mejores, aca solo puedes dibujar) y aparece en instagram algo de ""tu historia"" que es lo mismo de Snapchat! Todo bien, re buena pero, lo de que dura 24 hrs es RE COPIA de Snapchat, y no me gusta que sea copia porque esa es la gracia de Snapchat, que solo se vea algo 24 hrs, y si Instagram se copia, Snapchat no tiene gracia. SAQUEN TODO ESO PLIS!! ESTO ES PLAGIO. LE DARIA 5 ESTRELLAS SI NO TUVIESE TODO ESO.

**es-com.instagram.android-2-115**

La actualización...

Está bien, pero con la nueva actualización sólo salen las publicaciones mas relevantes y desordenadas cronológicamente, le falta ser un poco más intuitiva, gusta porque medio mundo está en Instagram, está de moda, no porque sea buena red social. Pd: A diferencia de otras redes sociales, instagram sólo es compatible con móviles y tableta, no con ordenadores. Y algo que tengo mucho en cuenta, solo se puede utilizar en vertical, y eso es incómodo. Las 2 estrellas con tantos errores son merecidas.

**es-com.instagram.android-4-116**

Puede mejorar

La app en sí está bien, pero podría ser mejor. Por ejemplo si me meto a ver una foto de alguien que sigo, me molesta tener que salir de esa foto y para ver la siguiente tener que meterme, y con la siguiente lo mismo. Sería mejor y mas sencillo poder pasarlas de lado o hacia arriba y abajo.... luego tiene un fallo de comunicacion o algo con flickr. Cuando

subo una foto y la quiero compartir directame te en facebook, twitter y flickr, se publica en todos sitios menos en flickr. Y tengo bien enlazada la cuenta

**es-com.instagram.android-4-117**

Mejoren una sola cosa...agreguen una opción tal cual como agregaron Boomerang o Layout

Agreguen en las opciones de edición de imágenes una opción similar a la app ""Instagrid"", así estaría perfecta ya que muchos usuarios de Instagram que conozco han descargado Instagrid para colocar en su cuenta pósters de 3×3, 3×4 y 3×5, espero este tip sirva para mejorar la app, ya que estos posters se ha viralizado su uso, Instagram por si es bastante buena, pero le falta actualizarse en cuanto a estas pequeñas tendencias, corrigen eso y tendrán las 5 estrellas!

**es-com.instagram.android-5-118**

Me ha encantado

Me encanta esta red social y encima la actualización que an hecho es fantástica seguro que le gusta a todo el mundo y por eso yo le doy 5 estrellas.....E leído los comentarios y a muchas personas le falla la nueva actualización. A mi no me falla nada ?? les doy la enhorabuena a los creadores de instagram y por favor no cambien la nueva actualización.

**es-com.instagram.android-5-119**

Esta muy buena

Esta muy buena la app instalensela , lo que más me gusta es q conoces a más personas de todos los países y te enteras de los famosos osea que están haciendo o grabando películas , vídeos o músicas . La recomiendo!

**es-com.instagram.android-5-120**

Problema

Hay un problema que tienen que arreglar con la nueva actualización sobre lo vídeos que uno comparte los audios no le sirve bien arreglen eso por favor para que se pueda escuchar bien cuando uno publica un vídeo y otra cosa que las letras sean mas pequeña cuando uno quiera escribir algo esta muy exagerado lo grande que sale las letras.

## **2.2. GAMES**

**es-com.king.candycrushsaga-1-121**

Es estúpido que cuando pasas el nivel 35 te pida pagar con dinero real o esperar 48 horas para seguir JUGANDO

Es un maldito juego, al pasar el lvl 35 te pide pagar con dinero real o te esperas 48 horas para seguir jugando... Que idiotas !!!! APARTE las 48 hrs es con el juego abierto porque si lo cierras no te cuenta el tiempo de espera...vayanse a la mierda de pana..que clase de



conformidad el de la población hoy en día. Que ven este tipo de reglas en un juego y aún así lo obvian porque simplemente está de ""moda"" .. zas

#### **es-com.king.candycrushsaga-1-122**

Aburrido

Voy en el nivel 1181, hasta el mil cien, mas o menos, me lo pasaba genial, había niveles muy difíciles, pero es la razón del juego, que nos esforcemos por solucionarlo ,pero ahora se ha vuelto aburrido, no da juego, te pasas la partida, haciendo el único movimiento que te da, así, una y otra y otra.Es que se les acabó, la imaginación?Me dan ganas de desinstalarlo.Vuelvo a quejarme!cada vez que hago una actualización, se cierra, perdiendo vidas, se jactan, de tenerlo en tantos países,como funcione, igual ...

#### **es-com.king.candycrushsaga-1-123**

Mala muy mala

Desde la última actualización se tilda y se sale el juego. Ademas cuándo estaba a un dulce de ganar en el nivel 389, compre un pack de \$119 pesos mexicanos para pagar los 5 tiros extras y el desgraciado juego se atoró y tuve que cerrar todo, y perdí la jugada ahora que lo juego no me da buenas jugadas, y se me acaban las jugadas y me pide gastar lingotes para 5 jugadas es un ROBOOOO!!!! Mejoren eso y doy 5 estrellas. (DESDE QUE SALIÓ EL JUEGO HE SIDO SU SEGUIDORA Y JUGADORA)

#### **es-com.king.candycrushsaga-2-124**

Pasar los niveles

Hay algunos difíciles y me llego a aburrir. Tengo q dejarlo sin jugar x un mes xa q se pueda pasar esos niveles....soy fan del candy pero puedo llegar a aburrirme si pasan días q no se desbloquea un nivel esperando q alguien lo haga o no hay blooster q me ayere a pasar o a desbloquear sin pedir a amigos. Muchos amigos están jugando al pokemon go....deberías reestructurar el juego xq sino ....no hay fan del candy

#### **es-com.king.candycrushsaga-2-125**

Fraude

Antes me gustaba pero no se pq me borro mi record hiba por mil y pico y me mando a empezar desde 0. Voy a bajar la aplicación de nuevo a ver si vuelvo en donde estaba si no la borraré, no es justo. La borre, no es justo q después de tanto tiempo jugando me borren mi nivel alcanzado.

#### **es-com.king.candycrushsaga-4-126**

Me gusta el juego

El juego no es Malo.. Son ustedes que no saben jugar.. El que dijo lo de Odus el búho estas pegado ahí porque tienes que ver abajo que dulces son a los que no le puedes llegar al final porque si no tumbas a Odus ... Yo voy por el ciento y algo y me da EXCELENTE EL JUEGO Así que no se QUEJEN

**es-com.king.candycrushsaga-4-127**

Esta fallando

El juego me encanta , pero cada vez falla mas . Se tilda y hay q reiniciar , en medio de un juego se corta y regresa a la pagina principal del telefono y lo peor es q desaparecen todos los premios q conseguiste y para desbloquear los capitulos tardas un monton porque no te toma lo q te envian . Una lástima porque perdes e l interés

**es-com.king.candycrushsaga-4-128**

A mi no me da ningun problema. Me entretiene y es muy bueno.... ademas de ello, es un juego que te hace pensar mucho y dinamico. Pensandolo desde otra perspectiva te ayuda a mejorar tu agilidad en la vista. Le daria las 5 estrellas sino fuese porque en veces se me pega. Pero en el resto del juego es muy bueno... disfrutento!!

**es-com.king.candycrushsaga-5-129**

Fallos constantes

Voy por el nivel 1507. Está muy bien; pero no se porqué, cada cierto tiempo al entrar he perdido todos los comodines acumulados. Y tenía treinta y tantos de cada; incluidas manos y chupachups. Y hoy en la última vez; he perdido también la opción esa de obtener un comodín cada día entrando en la app todos los días hasta 9 consecutivos. ¿Por qué pasa eso? Perder comodines sin saber porqué, después de lo que cuesta conseguirlos día a día. ¿Quieren que terminemos hartándonos de tantos fallos?

**es-com.king.candycrushsaga-5-130**

Es el mejor juego del mundo!

No estoy exagerando... Pero para los que dice que se tilda se abre se sierra pues jodansee tendrán un celular choto una tablr t choticima o peor rota o tendrán muchas aplicaciones pero esta es la mejor y si les pasa eso pues jodance y no digan después que él juego es una mierda y le dan una ?

**es-com.rovio.angrybirds-1-131**

Buen juego pero...

Son tontos a caso por poner el botón de pausa al lado de ese pájaro eléctrico? La mayoría los guardo para ocasiones especiales pero siempre al querer reiniciar apreto pausa y se pone el estúpido pájaro de mierda, como se les ocurre poner el pájaro al lado de ese nivel? Para variar cuando cierro el juego y vuelvo tengo menos pájaros a pesar de que no lo haya ocupado, es que hay que ser tonto para poner ese botón al lado del otro.

**es-com.rovio.angrybirds-1-132**

Demasiada publicidad

No creo que sea necesario hacer tan tedioso este juego, poner la publicidad en los últimos lanzamientos me parece legítimo pero hacerlo y a este nivel en el primero de la saga me parece demasiado. Lo he instalado después de varios años de haberlo jugado y ahora me

da miedo dejárselo a mi hija para que juegue un rato, después de un par de partidas lo he desinstalado.

**es-com.rovio.angrybirds-2-133**

Es bueno pero...

Es bueno pero, o sea pagar para que quiten los avisos, en eso se excidieron 2. Antes el mighty eagle se desbloqueaba por pasar sus 3 niveles y ahora tambien tienes que pagar para usarlo, espero que lo arreglen.

**es-com.rovio.angrybirds-2-134**

Lo tenia antes en mi tablet y servía ahora lo tengo en mi Xperia y no me deja jugarlo, lo pongo en un nivel y se mueven los angry birds pero no me deja jugarlo

**es-com.rovio.angrybirds-2-135**

Muy mal!!!

Todo esta bien pero desde la última actualización no lo puedo abrir si no estoy conectado al internet y se supone en que si se debería de poder jugar sin conexión..

**es-com.rovio.angrybirds-4-136**

Mmm...

No soy muy fan de estos juegos pero mmm esta bueno nada más tiene buena jugabilidad y gente si les molesta la publicidad desactiven su WiFi y listo

**es-com.rovio.angrybirds-4-137**

Esta bueno y entretenido, es el clásico Angry Birds, y ya desbloquee el rey cerdo dorado, para hacerlo tienes que tener tres estrellas en Poached eggs hasta Birdday parthy

**es-com.rovio.angrybirds-5-138**

Sin duda descargenlo

El juego para la cantidad de mundos q tiene no pesa tanto así y se mueve rápido no se me pega y tampoco se sale y si te llaman o te suena la alarma no se sale por lo menos a mi me funciona bien pero lo único malo es los trucos y poderes se compran con dinero de verdad y yo no tengo dolares soy Venezolano y si conocen al país no ganamos dolares así que fuera fino si al menos los trucos pequeños se compraran con huevos de oro que uno se consigue escondido en algunos mundos y paginas de los mundos

**es-com.rovio.angrybirds-5-139**

Muy buena

A mí me parece que en la aplicación pusieran a red .el que tiene poder. junto con bomb electrificadora para que cuando aya un nivel complicado tambien utilizarlo. Tengo un solo problema , que cuando quiero jugar o entrar a la aplicación . Aparece el dibujo de angry birds por mucho tiempo y se me traba el celu de todas maneras es una aplicación excelente . Y la peli es excelente y coincide . Esos anuncios hay que decirles adios .



ahí que ser un completo estúpido, que no sabe de calidad literaria para, llegar siquiera a creer que este libro es bueno, historia acelerada, personajes sin fuerza, en si la mayor perdida de mi tiempo en este genero. y he leído la obras del marques de sade. el erotismo es algo a lo que no le veo tanta relevancia como para usarlo como medio mediocre para darle peso a la paginas

**es-Cincuenta sombras de Grey-1-147**

Sobrevalorado

El argumento me parece aburrido (por no decir que no tiene) y no sé a que viene tanta emoción por un libro tan mediocre. Además de que quitando el sexo el libro es machista a más no poder, y creo que se puede comparar a la saga crepúsculo. Y al que le moleste, que lea cosas mejores . En fin, una decepción para mi.

**es-Cincuenta sombras de Grey-2-148**

Puro sexo

Cual es el fin, lo único que se aprecia es el sexo, no mne refiero a que se aprecia, si no lo que gira en torno al libro, es solo sexo, me sorprende que la gente por lo único que lo lea sea por eso pueden decir que es por otra, pero realmente lo único que quieren es leerlo por esta razón, no me refiero a todo@s, si no a la mayoría.

**es-Cincuenta sombras de Grey-2-149**

Muy malo

Q diga es lo mejor k se ha leído nunca me da ganas d llorar q injusticia por dios jajajaja con los buenos escritores y narradores k existen. Si mi hija narra los hechos mejor q esta tia ayyy en q mundo vivimos

**es-Cincuenta sombras de Grey-2-150**

Infumable

Tirando a malo. Lo único que me duele es que después de leer el libro hay mucha gente que opina pero no saben ni escribir sin hacer faltas horribles de escritura. ¿ Será que escuchan en vez de leer? Para CATETOS

**es-Cincuenta sombras de Grey-4-151**

Luces y sombras un lugar de reencuentro interior

A simple vista..puede parecer una novela exclusivamente femenina, para aliviar y conectar con las pasiones y deseos innato a aletargados. Pero encierra mucho mas...remueve interiormente una lucha consigo misma y ayuda a comprender a la persona que tienes enfrente, con la simple observación y contemplacion... Todos tenemos luces y sombras, y están esperando a que las conozcamos. Recomendando la novela a todas las personas, que necesitan reencontrarse y soñar....

**es-Cincuenta sombras de Grey-5-152**

Debéis leerlo!!!

Es fascinante. Algunos momentos me dejaron disgustada e incluso enfadado con Ana por no saber defenderse como Dios manda. Por eso el final me gusto bastante. Faltaba una buena bofetada antes de irse. Pero por cierto el segundo libro y el tercero me gustaron más. Tenéis que leer la trilogía entera para entender las ideas de este primer libro.

**es-La chica del tren-1-153**

Ritmo lento y contenido indiferente

Me lo recomendó mi madre aún sin leerlo ya que se lo habían recomendado a ella. Tuve la desgracia de perder mi tiempo leyéndolo antes que ella. Es un libro lento y pesado, con un título engañoso, que te cuenta una historia que te deja completamente indiferente... Si esperas un thriller o algo sobrenatural sigue buscando porque en este libro no lo encontrarás.

**es-La chica del tren-1-154**

Predecible

El principio se hace pesado y ya las ultimas 100 pag se va desarrollando la trama y al final sabes quien fue y como ocurrio,es predecible y adivinas.me esperaba otra cosa mas impactante un bien 6

**es-La chica del tren-2-155**

Sobre valorado

La publicidad que se le ha realizado a éste libro, pasa de ser rimbombante!. No niego que inicia con una puesta en escena interesante, pero para ser un thriller, no se incluyeron piezas para ello; es más, casi todo el transcurso del libro narró la vida de parejas, problemas matrimoniales, y una vida destrozada por parte de la protagonista. El misterio, la desaparición estuvo casi a un plano terciario.. No hay abogados, no hay detectives (si aparecen pero no aportan absolutamente nada a la historia)

**es-La chica del tren-2-156**

Lioso!!

Tiene algo de intriga...y es lo que te hace terminar de leerlo...pero no engancha...esperas todo el tiempo que tenga algun giro inesperado....que nunca llega!...??

**es-La chica del tren-4-157**

Me asombró

Narra muy bien la psique de los/as protagonistas y el trato del entorno a Ellos/as. Me asombró el final. Los primeros capítulos se hacen un poco más pesados, pero luego se lee rápido y engancha. No es lectura para todo el mundo, hace pensar.

**es-La chica del tren-4-158**

Ameno y de calidad.

Una telaraña de tramas paralelas que nos enfrentan a nuestros miedos sociales, nuestras relaciones personales, en las que la vida anhelada vista como espectadora puede ser una agonía no tan idílica en la que nada es lo que parece.

**es-La chica del tren-4-159**

Entretiene

Mucha publicidad para una historia más. Los personajes son bien elaborados, la trama logra atraparte y la narrativa es buena. Pero nada especial. A mitad libro ya deducir es criminal. Me alegra haberla comprado por acá y no impresa.

**es-La chica del tren-5-160**

La chica del tren

Tube la oportunidad de comprar este libro y la verdad que no me arrepiento. Es lo mejor que leí, la manera en la que el autor narra la historia es buenisima. Tiene mucho suspenso y te hace pensar y sacar tus propias conclusiones pero al final de la historia todo lo que vos pensabas cambia. Paso a ser mi novela favorita. Les recomiendo que la lean, van a quedar encantados.

## 2.4. FILMS

**es-Avatar-1-161**

Muy cara.

¿De verdad este precio por una película que ni es HD, con lo antigua que es, y que ni siquiera es la versión extendida? Si queréis que la gente pague, poned precios razonables, y no esto. Iba a comprarla para empezar a coleccionar vídeos en google play movies, pero visto lo visto he vuelto a cambiar de opinión. No quiero pagar estos precios. Una cosa es que me intenten engañar, y otra muy distinta es que me deje. Gracias por ponerme claras las cosas.

**es-Avatar-1-162**

Pocahontas del espacio

Fue lo que pensé cuando anunciaban de qué iba la peli. Con los años al fin la vi... y pues eso... pocahontas espacial. Los extras donde se muestran las técnicas para hacer la peli son los que valen la pena... la película, para niños.

**es-Avatar-2-163**

el precio exagerado

lo dicho el precio es exagerado. mejor comprarla en formato fisico ke sale mas barato. se supone ke las descargas deberian ser mas baratas esto es un sin sentido

**es-Avatar-2-164**

carisimo

De momento el consumidor de a pie, no puede presumir de ser un habitual en compras de películas online, por el precio excesivo de las visualizaciones.

**es-Avatar-4-165**

Preciosa

Es muy bonita recomiendo que la vean, yo al principio no queria verla porque pensaba que era futurista y mala pero es genial muy buena veanla. P.D. Son muy caras bajen los precios de las películas.

**es-Avatar-5-166**

A lo grande

James Cameron, hace las cosas a lo grande. Ya lo hizo con titanic y con terminator. Esta película es la mejor que ha hecho en toda su carrera profesional en el mundo del cine.

**es-Avatar-5-167**

Obra maestra

Simplemente una obra maestra, la mejor película de la historia que jamás se haya creado y que jamás se creará, James Cameron simplemente es un genio una obra digna de ver y de admirar simplemente PERFECTA

**es-Frozen\_ El Reino del Hielo-1-168**

Muy sobrevalorada

No entiendo la fama que ha cogido esta película, cuando es una de las más mediocres y peores que he visto de Disney. Odio que la gente sobrevalore algo que no tiene valor. Desde el primer momento no quise verla, así que ni siquiera gasté mi dinero en ir al cine. Como todo el mundo estaba diciendo lo buena que era, pues al final la vi, y confirmé que era tan mala como me parecía.

**es-Frozen\_ El Reino del Hielo-2-169**

Ya estoy hartito

La 1a vez que la ví me gustó,pero ya un billón de veces como que ya no. A parte ya han puesto juegos feos que ya quiero poner una demanda,y en mi anterior colegio y ahora está mi hermana pequeña,a los niños les ponen Bob Esponja o superhéroes y a las niñas Frozen y ya no puedo más y todo el rato lo mismo aburre. Es por eso que le pongo 2 estrellas, no le pongo 1 estrella porque no está mal que si no...

**es-Frozen\_ El Reino del Hielo-4-170**

Sobre valorada

Es una película muy buena, y admito que me gustó mucho, pero aún así, pienso que está sobre valorada. Tiene más éxito del que se merece. Cada uno tiene su opinión y yo opino esto. ¡Saludos!



**es-Frozen\_ El Reino del Hielo-4-171**

Fantastica

Disney tiene una película genial, canciones fantásticas, muy tierna y divertida. La recomiendo a toda clase de públicos. Se han superado, con esta película. Me encantaaaa!!!

**es-Frozen\_ El Reino del Hielo-5-172**

Me encantó esta película!

Es una nueva obra maestra de los estudios Walt Disney, mostrando al público la impresionante amistad entre dos hermanas que por un accidente de una de las dos, las condena a una separación eterna. Es una pasada de película, los efectos especiales muy mágicos y bien trabajados, Arendelle, impresionante e imponente, el castillo me encantó, Olaf súper tierno (creo que se me pegó lo de los abrazitos calentitos jeje), Elsa me conmovió de cómo intentaba proteger a su hermana, y cuando huye congelando el fiordo...impresionante...de mis escenas favoritas. Las canciones muy bonitas, y yo me enrolló más que las persianas... Muy bonita la película y se merece todos los premios cinematográficos que puedan existir en el mundo entero!! Y confieso: Elsa, Anna y Olaf son mis favoritos (especialmente Elsa). Le doy un diez y 5 estrellas

**es-Frozen\_ El Reino del Hielo-5-173**

Me ha encantado

Es genial y la he visto en el cine no me estrañan los comentarios y q la peli alla ganado a la megor peli de dos oscars animacion y musica y porcierto Jose tu siempre dices eso y los padres de Ana y Elsa siguen muertos asique no te pongas chulito y no digas vovadas y Enredados también está genial tu dices que es malo que todo acabe bien por que tu eres perfecto aque si y si no es así porque lo dices a ti te gustaría que si un cuento tratase de ti acabara mal no se tu pero a mi no pero como tu estas mal

**es-Gravity-1-174**

Es gratis pero....

La verdad es que no es buena. Aunque sea gratis

**es-Gravity-1-175**

Un tostón

Aburrida, bonito intento de hacerte desesperar. Lo único que consigue es hacerte bostezar

**es-Gravity-2-176**

Buff, muy mala

Lo único que se salva es la fotografía y ver la tierra y las estaciones espaciales. Hay cosas que no tienen sentido, no se, parece una americanada, pero en el espacio. Si has visto interstellar, no te recomiendo que veas gravity.

**es-Gravity-2-177**

Esta película no es para tanto ,ha habido muchisimas mejores y mas emocionantes como LA TEORIA DEL TODO es muchisimo mas perfecta y elaborada no entiendo como esta peli esta tan arriba

**es-Gravity-4-178**

Buenos efectos especiales

Y una historia mejorable, la película bien se podría llamar la increíble potra de la astronauta más torpe de la historia, pero lo resumieron en Gravity LOL ??. De todas formas un 10 para Google por regalar a los compradores de un Nexus ??.

**es-Gravity-4-179**

Peli para ver en 3D por efectos. No esta mal la historia.

La peli, es un alarde de Fotografía y efectos especiales. Pena que la vi en 2D en casa. Digna de ir al cine y en 3D ! La historia ,esta bien.

**es-Gravity-5-180**

Impresionante, y en 3D es insuperable.

Los efectos especiales, el desarrollo, el ritmo que tiene la película, los sonidos te envuelven, la música aún más. En 3D en tele 3D se aprecian cosas que en una televisión normal no se ve, como son las chispas del fuego, las gotas de agua, el humo, el tornillo, el reflejo al ver el planeta Tierra se aprecia en 3d y que en 2d no sabes ni que está. Mi favorita sin duda. Me ha fascinado.

## **2.5. MUSIC**

**es-AC\_DC Back in Black-1-181**

Todos los discos de acdc son super aburridos buaah

**es-AC\_DC Back in Black-4-182**

Wow!! Me gusto

Todas las rolas esta padrisimas buenos tiempos aquellos

**es-AC\_DC Back in Black-4-183**

Acdc

Mi grupo favorito acdc lo mejor que pudo haber existido lml

**es-AC\_DC Back in Black-5-184**

El mejor disco de hard-rock de todos los tiempos.

La situación de la banda (por la muerte de su vocalista y contratación de uno nuevo) y la época en que sucede la grabación de esta obra maestra, hace de este album de AC/DC un parteaguas en el género del hard-rock, y crea un referente en la historia de la música.

Temas como Hells Bells, Shoot to thrill, Back in Black y You shook me all night long sobresalen de la lista de temas de este disco.

#### **es-AC\_DC Back in Black-5-185**

La mejor banda del mundo

No cabe duda que con el pasar de los años AC/DC se convirtió en una de las bandas más importantes e influyentes del rock internacional.

#### **es-Adele25-1-186**

MUY BUEN ALBUM ??

La verdad es que merece un excellent Adele...por su voz trascendental y un don con el que nació. Pero esa es la cuestión, ella no debe esforzarse por cantar bien sino por dar un buen trabajo y calidad a sus composiciones. No voy a negar que oírla es algo maravilloso, algo indiscutible. Pero su composición es pobre. Escuchar sus 3 álbumes es escuchar repetidas veces una misma canción, no va adquiriendo nuevas habilidades y madurando en la música. Sigue estancada en una misma situación de privilegio de voz.

#### **es-Adele25-1-187**

Sepan esto...

No se porque dicen que es un album increible solo estamos volviendo a escuchar la tipica musica de Adele este album solo es igual que 19 y 21. Si acaso solo una cancion es buena pero no deja de parecerse alas anteriores pero no podemos evitar bostesar con el resto de album no tiene ritmo no son diferentes ni nada no es algo nuevo

#### **es-Adele25-1-188**

Contiene la misma estructura musical de sus demás discos, con excepción de los toques digitales justificados por la falta de innovación de la composición sonora. Además de presentar plagio y un monotonismo análogo a su discografía total , con ayuda extra en el mejoramiento de la voz por métodos digitales para evitar sonar desafinada ( como varios )

#### **es-Adele25-1-189**

Que ingenuos

Podrian ahorrarse tanto dinero convirtiendo el video en audio MP3 y luego descargarlo en el telefono.Es piratear pero se escucha perfecto y ademas es gratuito.Pueden hacer eso con todas las canciones que quieran y fin.Realmente me da mucha gracia todo aquel que gaste en esto

#### **es-Adele25-2-190**

X

Le tenía más expectativas al álbum por toda la controversia que creó y más tratándose de Adele, pero solo es otro álbum más. Adele es de las mejores artistas del momento, pero muy repetitiva.

**es-Adele25-2-191**

Falló la descarga

Me encanta Adele y el disco en si me gusto mucho pero ahora que intente descargarlo solo se hizo el cobro pero me aparece error al querer reproducir el álbum

**es-Adele25-2-192**

La verdad es que el album tiene muy buena musica, pero es de esos albums que solo lo escuchas una sola vez y lo guardas.

**es-Adele25-2-193**

una voz inigualable la de adele. en este album despliega todosu talento vocal e interpretativo, en altos y bajos melodias del alma

**es-Adele25-2-194**

Bah, la mayoría no soy buenas

Y sus vocales siempre son iguales, son muy bueno pero no son los mejores

**es-Adele25-4-195**

Le doy 4 estrellas

El álbum es muy bueno (y lo amo) pero no me gusto el orden de algunas canciones: water under the bridge y river lea son las dos canciones movidas que tiene el álbum y las pone a las dos juntas, me hubiese gustado que el tracklist sea - water under the bridge - million years ago - river lea - love in the dark. Al fin, muy buen álbum, me gustaron todas las canciones. Mis 3 canciones favoritas son: hello - i miss you - all i ask

**es-Adele25-4-196**

Falta letras y código

Me han regalado el disco físico de Adele y no veo bien que para tenerlo en formato digital haya que comprarlo otra vez. Ya que es más caro en disco debería haber un código que pudiéramos escuchar las canciones gratis de forma digital. También muchas fotos de ella y no vienen las letras.

**es-Adele25-4-197**

Me re gusto no le doy 5 estrellas por que todas las canciones no me gustan

Esta re bueno el album siempre me gusto adele pero ahora me encanta,...mi cancion favorita de este albun es hello....

**es-Adele25-5-198**

El Mejor hasta la Fecha

Para hablar de Adele, hay que empezar por decir que es maravillosa. Este nuevo disco ha sido el gran pasó que necesita un cantante tan famoso como para empezar a afirmar que tan fenomenal es. Ha hecho de todo con este álbum, sin duda alguna se ha renovado y eso lo sienten todos los que la seguimos desde hace ya rato. Realmente se esforzó en cada

letra y parecía que intentó encontrar esa nostalgia. La verdad la podría escuchar por horas y horas. Buen trabajo, lírica redonda y producción de primera. Excelente

**es-Adele25-5-199**

**AMO A ADELE**

Simplemente es perfecto, genial, hermoso, esta mujer se pasa ya, es muy perfecta en todos los sentidos &3 4 años esperando que saque este glorioso album que es la perfección como ella, la amo y la amo y que siga con su humildad y cariño que siempre tiene, además, ella hace musica para él oído no para los ojos como hacen miley cyrus o algunas de esas locas drogadictas que se andan desnudndo, Adele lo que hace es musica de la buena y listo. La amo

**es-Adele25-5-200**

**HERMOSO**

Sin dudas un disco maravilloso, me atrevo a decir que es uno de los mejores de la historia de la musica, todas las canciones tienen excelente letra y una impecable interpretación. Muy sentimental, toca el centro de las emociones de todo aquel que lo escuche. Supera al disco 21 pues Adele se muestra mas madura y consagrada. Simplemente me encanta el disco y espero que Adele consiga muchos éxitos con él. Estoy feliz de tener este disco que sin dudas es una joya de la musica mundial. ¡¡¡ Felicidades Adele!!!



## **Appendix G. ANNOTATED CORPORA**

### **G.1. ANNOTATED TRAINING CORPUS**

The full annotated training corpus for each of the annotation experiments is contained in the CD attached.

### **G.1. ANNOTATED LARGER CORPUS**

The full annotated larger corpus is contained in the CD attached. It requires the installation of the UAM Corpus Tool.